



Europäisches Patentamt
European Patent Office
Office européen des brevets



(11) Publication number: **0 251 446 B1**

(12)

EUROPEAN PATENT SPECIFICATION

- (43) Date of publication of patent specification: **28.12.94** (51) Int. Cl.⁵: **C12N 15/00, C12N 9/54, C12N 1/00**
- (21) Application number: **87303761.8**
- (22) Date of filing: **28.04.87**

- (54) **Non-human Carbonyl hydrolase mutants, DNA sequences and vectors encoding same and hosts transformed with said vectors.**

- (30) Priority: **30.04.86 US 858594**
06.04.87 US 35652
- (43) Date of publication of application:
07.01.88 Bulletin 88/01
- (45) Publication of the grant of the patent:
28.12.94 Bulletin 94/52
- (84) Designated Contracting States:
AT BE CH DE ES FR GB GR IT LI LU NL SE
- (56) References cited:
EP-A- 0 130 756
WO-A-87/04461
WO-A-87/05050

**ABSTRACTS OF THE 190TH AMERICAN
CHEMICAL SOCIETY NATIONAL MEETING,
vol. 190,1985, page 23, no. 47; R.R. BOTT et
al.: "Protein engineering of subtilisin"**

- (73) Proprietor: **GENENCOR INTERNATIONAL, INC.**
180 Kimball Way
South San Francisco, CA 94080 (US)
- (72) Inventor: **Wells, James Allen**
64 Otay Avenue
San Mateo
CA 94403 (US)
Inventor: **Cunningham, Brian C.**
24 Olive Avenue
Piedmont
CA 94611 (US)
Inventor: **Caldwell, Robert Mark**
1828 Broadway
No.101
San Francisco
Ca 94109 (US)
Inventor: **Bott, Richard Ray**
3032 Hillside drive
Burlingame
CA 94010 (US)

Note: Within nine months from the publication of the mention of the grant of the European patent, any person may give notice to the European Patent Office of opposition to the European patent granted. Notice of opposition shall be filed in a written reasoned statement. It shall not be deemed to have been filed until the opposition fee has been paid (Art. 99(1) European patent convention).

EP 0 251 446 B1

JOURNAL OF CELLULAR BIOCHEMISTRY SUPPL., vol. 0, no. 10, part A, 1986, page 271, no. E101, SYMPOSIUM ON PROTEASES IN BIOLOGICAL CONTROL AND BIOTECHNOLOGY, 15th ANNUAL UCLA, MEETING ON MOLECULAR AND CELLULAR BIOLOGY, Los Angeles, CA, 9th-15th February 1986; P. BRYAN et al.: "Protein engineering of subtilisin-proteases of enhanced stability"

WORLD BIOTECH. REPORT, vol. 2, 1985, pages 51-59, Online Publications, Pinner, GB; R. BOTT: "Modeling & crystallographic analysis of site-specific mutants of subtilisin"

JOURNAL OF CELLULAR BIOCHEMISTRY SUPPL., vol. 0, no. 11, part C, 1987, page 200, no. N024, New York, US; D.A. ESTELL et al.: "Tailoring enzymatic properties through multiple mutations"

PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCE USA, vol. 84, March 1987, pages 1219-1223, Washington, D.C., US; J.A. WELLS et al.: "Designing substrate specificity by protein engineering of electrostatic interactions"

BIOCHEMISTRY, vol. 26, no. 8, April 1987, pages 2077-2082, American Chemical Society, Washington, D.C., US; M.W. PANTOLIANO et al.: "Protein engineering of subtilisin BPN': enhanced stabilization through the introduction of two cysteines to form a disulfide bond"

PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCE USA, vol. 83, June 1986, pages 3743-3745, Washington, D.C., US; P. BRYAN et al.: "Site-directed mutagenesis and the role of the oxyanion hole in subtilisin"

NATURE, vol. 318, 28th November 1985, pages 375-376, London, GB; P.G. THOMAS et al.: "Tailoring the pH dependence of enzyme catalysis using protein engineering"

JOURNAL OF BACTERIOLOGY, vol. 158, no. 2, May 1984, pages 411-418, American Society for Microbiology, Washington, D.C., US; M.L. STAHL et al.: "Replacement of the Bacillus subtilis subtilisin structural gene with an in vitro-derived deletion mutation"

Inventor: Estell, David Aaron
250 Diablo Avenue
Mountain View
CA 94043 (US)

Inventor: Power, Scott Douglas
732 Olive Court
San Bruno
CA 94066 (US)

⑦ Representative: Armitage, Ian Michael et al
MEWBURN ELLIS
York House
23 Kingsway
London WC2B 6HP (GB)

NUCLEIC ACIDS RESEARCH, vol. 11, no. 22,
November 1983, pages 7911-7925, IRL Press
Ltd, Cambridge, GB; J.A. WELLS et al.: "Clon-
ing, sequencing, and secretion of Bacillus
amyloliquefaciens subtilisin in Bacillus sub-
tilis"

Description

The recent development of various *in vitro* techniques to manipulate the DNA sequences encoding naturally-occurring polypeptides as well as recent developments in the chemical synthesis of relatively short sequences of single and double stranded DNA has resulted in the speculation that such techniques can be used to modify enzymes to improve some functional property in a predictable way. Ulmer, K.M. (1983) *Science* 219, 666-671. The only working example disclosed therein is the substitution of a single amino acid within the active site of tyrosyl-tRNA synthetase (Cys35→Ser) which lead to a reduction in enzymatic activity. See Winter, G., et al. (1982) *Nature* 299, 756-758; and Wilkinson, A.J., et al. (1983) *Biochemistry* 22, 3581-3586 (Cys35→Gly mutation also resulted in decreased activity).

When the same t-RNA synthetase was modified by substituting a different amino acid residue within the active site with two different amino acids, one of the mutants (Thr51→Ala) reportedly demonstrated a predicted moderate increase in *kcat*/*Km* whereas a second mutant (Thr51→Pro) demonstrated a massive increase in *kcat*/*Km* which could not be explained with certainty. Wilkinson, A.H., et al. (1984) *Nature* 307, 187-188.

Another reported example of a single substitution of an amino acid residue is the substitution of cysteine for isoleucine at the third residue of T4 lysozyme. Perry, L.J., et al. (1984) *Science* 226, 555-557. The resultant mutant lysozyme was mildly oxidized to form a disulfide bond between the new cysteine residue at position 3 and the native cysteine at position 97. This crosslinked mutant was initially described by the author as being enzymatically identical to, but more thermally stable than, the wild type enzyme. However, in a "Note Added in Proof", the author indicated that the enhanced stability observed was probably due to a chemical modification of cysteine at residue 54 since the mutant lysozyme with a free thiol at Cys54 has a thermal stability identical to the wild type lysozyme.

Similarly, a modified dihydrofolate reductase from *E. coli* has been reported to be modified by similar methods to introduce a cysteine which could be cross linked with a naturally-occurring cysteine in the reductase. Villafranca, D.E., et al. (1983) *Science* 222, 782-788. The author indicates that this mutant is fully reactive in the reduced state but has significantly diminished activity in the oxidized state. In addition, two other substitutions of specific amino acid residues are reported which resulted in mutants which had diminished or no activity.

EPO Publication No. 0130756 discloses the substitution of specific residues within *B. amyloliquefaciens* subtilisin with specific amino acids. Thus, Met222 has been substituted with all 19 other amino acids, Gly166 with 9 different amino acids and Gly169 with Ala and Ser.

As set forth below, several laboratories have also reported the use of site directed mutagenesis to produce the mutation of more than one amino acid residue within a polypeptide.

The amino-terminal region of the signal peptide of the prolipoprotein of the *E. coli* outer membrane was stated to be altered by the substitution or deletion of residues 2 and 3 to produce a charge change in that region of the polypeptide. Inouye, S., et al. (1982) *Proc. Nat. Acad. Sci. USA* 79, 3438-3441. The same laboratory also reported the substitution and deletion of amino acid residues 9 and 14 to determine the effects of such substitution on the hydrophobic region of the same signal sequence. Inouye, S., et al. (1984) *J. Biol. Chem.* 259, 3729-3733.

Double mutants in the active site of tyrosyl-t-RNA synthetase have also been reported. Carter, P.J., et al. (1984) *Cell* 38, 835-840. In this report, the improved affinity of the previously described Thr51→Pro mutant for ATP was probed by producing a second mutation in the active site of the enzyme. One of the double mutants, Gly35/Pro51, reportedly demonstrated an unexpected result in that it bound ATP in the transition state better than was expected from the two single mutants. Moreover, the author warns, at least for one double mutant, that it is not readily predictable how one substitution alters the effect caused by the other substitution and that care must be taken in interpreting such substitutions.

A mutant is disclosed in U.S. Patent No. 4,532,207, wherein a polyarginine tail was attached to the C-terminal residue of β -urogastrone by modifying the DNA sequence encoding the polypeptide. As disclosed, the polyarginine tail changed the electrophoretic mobility of the urogastrone-polyarginine hybrid permitting selective purification. The polyarginine was subsequently removed, according to the patentee, by a polyarginine specific exopeptidase to produce the purified urogastrone. Properly construed, this reference discloses hybrid polypeptides which do not constitute mutant polypeptides containing the substitution, insertion or deletion of one or more amino acids of a naturally occurring polypeptide.

Single and double mutants of rat pancreatic trypsin have also been reported. Craik, C.S., et al. (1985) *Science* 228, 291-297. As reported, glycine residues at positions 216 and 226 were replaced with alanine residues to produce three trypsin mutants (two single mutants and one double mutant). In the case of the single mutants, the authors stated expectation was to observe a differential effect on *Km*. They instead

reported a change in specificity (kcat/Km) which was primarily the result of a decrease in kcat. In contrast, the double mutant reportedly demonstrated a differential increase in Km for lysyl and arginyl substrates as compared to wild type trypsin but had virtually no catalytic activity.

The references discussed above are provided solely for their disclosure prior to the filing date of the instant case, and nothing herein is to be construed as an admission that the inventors are not entitled to antedate such disclosure by virtue of prior invention or priority based on earlier filed applications.

Based on the above references, however, it is apparent that the modification of the amino acid sequence of wild type enzymes often results in the decrease or destruction of biological activity.

Accordingly, it is an object herein to provide carbonyl hydrolase mutants which have at least one property which is different from the same property of the carbonyl hydrolase precursor from which the amino acid of said mutant is derived.

It is a further object to provide mutant DNA sequences encoding such carbonyl hydrolase mutants as well as expression vectors containing such mutant DNA sequences.

Still further, another object of the present invention is to provide host cells transformed with such vectors as well as host cells which are capable of expressing such mutants either intracellularly or extracellularly.

Summary of the Invention

The invention includes carbonyl hydrolase mutants, preferably having at least one property which is substantially different from the same property of the precursor non-human carbonyl hydrolase from which the amino acid sequence of the mutant is derived. These properties include oxidative stability, substrate, specificity catalytic activity, thermal stability, alkaline stability, pH activity profile and resistance to proteolytic degradation. The precursor carbonyl hydrolase may be naturally occurring carbonyl hydrolases or recombinant carbonyl hydrolases. The amino acid sequence of the carbonyl hydrolase mutant is derived by the substitution, deletion or insertion of one or more amino acids of the precursor carbonyl hydrolase amino acid sequence.

The invention also includes mutant DNA sequences encoding such carbonyl hydrolase mutants. Further the invention includes expression vectors containing such mutant DNA sequences as well as host cells transformed with such vectors which are capable of expressing said carbonyl hydrolase mutants.

Brief Description of the Drawings

Figure 1 shows the nucleotide sequence of the coding strand, correlated with the amino acid sequence of B. amyloliquefaciens subtilisin gene. Promoter (p) ribosome binding site (rbs) and termination (term) regions of the DNA sequence as well as sequences encoding the presequence (PRE) putative prosequence (PRO) and mature form (MAT) of the hydrolase are also shown.

Figure 2 is a schematic diagram showing the substrate binding cleft of subtilisin together with substrate.

Figure 3 is a stereo view of the S-1 binding subsite of B. amyloliquefaciens subtilisin showing a lysine P-1 substrate bound in the site in two different ways. Figure 3A shows Lysine P-1 substrate bound to form a salt bridge with a Glu at position 156. Figure 3B shows Lysine P-1 substrate bound to form a salt bridge with Glu at position 166.

Figure 4 is a schematic diagram of the active site of subtilisin Asp32, His64 and Ser221.

Figures 5A and 5B depict the amino acid sequence of subtilisin obtained from various sources. The residues directly beneath each residue of B. amyloliquefaciens subtilisin are equivalent residues which (1) can be mutated in a similar manner to that described for B. amyloliquefaciens subtilisin, or (2) can be used as a replacement amino acid residue in B. amyloliquefaciens subtilisin. Figure 5C depicts conserved residues of B. amyloliquefaciens subtilisin when compared to other subtilisin sequences.

Figures 6A and 6B depict the inactivation of the mutants Met222L and Met222Q when exposed to various organic oxidants.

Figure 7 depicts the ultraviolet spectrum of Met222F subtilisin and the difference spectrum generated after inactivation by diperdodecanoic acid (DPDA).

Figure 8 shows the pattern of cyanogen bromide digests of untreated and DPDA oxidized subtilisin Met222F on high resolution SDS-pyridine peptide gels.

Figure 9 depicts a map of the cyanogen bromide fragments of Fig. 8 and their alignment with the sequence of subtilisin Met222F.

Figure 10 depicts the construction of mutations between codons 45 and 50 of B. amyloliquefaciens subtilisin.

Figure 11 depicts the construction of mutations between codons 122 and 127 of B. amyloliquefaciens subtilisin.

Figure 12 depicts the effect of DPDA on the activity of subtilisin mutants at positions 50 and 124 in subtilisin Met222F.

5 Figure 13 depicts the construction of mutations at codon 166 of B. amyloliquefaciens subtilisin.

Figure 14 depicts the effect of hydrophobicity of the P-1 substrate side-chain on the kinetic parameters of wild-type B. amyloliquefaciens subtilisin.

Figure 15 depicts the effect of position 166 side-chain substitutions on P-I substrate specificity. Figure 15A shows position 166 mutant subtilisins containing non-branched alkyl and aromatic side-chain substitu-
10 tions arranged in order of increasing molecular volume. Figure 15B shows a series of mutant enzymes progressing through β - and γ -branched aliphatic side chain substitutions of increasing molecular volume.

Figure 16 depicts the effect of position 166 side-chain volume on log kcat/Km for various P-1 substrates.

Figure 17 shows the substrate specificity differences between Ile166 and wild-type (Gly166) B. amyloliquefaciens subtilisin against a series of aliphatic and aromatic substrates. Each bar represents the
15 difference in log kcat/Km for Ile166 minus wild-type (Gly166) subtilisin.

Figure 18 depicts the construction of mutations at codon 169 of B. amyloliquefaciens subtilisin.

Figure 19 depicts the construction of mutations at codon 104 of B. amyloliquefaciens subtilisin.

Figure 20 depicts the construction of mutations at codon 152 B. amyloliquefaciens subtilisin.

20 Figure 21 depicts the construction of single mutations at codon 156 and double mutations at codons 156 and 166 of B. amyloliquefaciens subtilisin.

Figure 22 depicts the construction of mutations at codon 217 for B. amyloliquefaciens subtilisin.

Figure 23 depicts the kcat/Km versus pH profile for mutations at codon 156 and 166 in B. amyloliquefaciens subtilisin.

25 Figure 23A depicts the kcat/Km versus pH profile for mutations at codon 156 and 166 in B. amyloliquefaciens subtilisin.

Figure 24 depicts the kcat/Km versus pH profile for mutations at codon 222 in B. amyloliquefaciens subtilisin.

Figure 25 depicts the constructing mutants at codons 94, 95 and 96.

30 Figures 26 and 27 depict substrate specificity of various wild type and mutant subtilisins for different substrates.

Figures 28 A, B, C and D depict the effect of charge in the P-1 binding sites due to substitutions at codon 156 and 166.

Figures 29 A and B are a stereoview of the P-1 binding site of subtilisin BPN' showing a lysine P-1
35 substrate bound in the site in two ways. In 29A, Lysine P-1 substrate is built to form a salt bridge with a Glu at codon 156. In 29B, Lysine P-1 substrate is built to form a salt bridge with Glu at codon 166.

Figure 30 demonstrates residual enzyme activity versus temperature curves for purified wild-type (Panel A), C22/C87 (Panel B) and C24/C87 (Panel C).

Figure 31 depicts the strategy for producing point mutations in the subtilisin coding sequence by
40 misincorporation of α -thioldeoxynucleotide triphosphates.

Figure 32 depicts the autolytic stability of purified wild type and mutant subtilisins 170E, 107V, 213R and 107V/213R at alkaline pH.

Figure 33 depicts the autolytic stability of purified wild type and mutant subtilisins V50, F50 and
45 F50/V107/R213 at alkaline pH.

Figure 34 depicts the strategy for constructing plasmids containing random cassette mutagenesis over
residues 197 through 228.

Figure 35 depicts the oligodeoxynucleotides used for random cassette mutagenesis over residues 197
through 228.

Figure 36 depicts the construction of mutants at codon 204.

50 Figure 37 depicts the oligodeoxynucleotides used for synthesizing mutants at codon 204.

Detailed Description

55 The inventors have discovered that various single and multiple in vitro mutations involving the substitution, deletion or insertion of one or more amino acids within a non-human carbonyl hydrolase amino acid sequence can confer advantageous properties to such mutants when compared to the non-mutated carbonyl hydrolase.

Specifically, *B. amyloliquefaciens* subtilisin, an alkaline bacterial protease, has been mutated by modifying the DNA encoding the subtilisin to encode the substitution of one or more amino acids at various amino acid residues within the mature form of the subtilisin molecule. These *in vitro* mutant subtilisins have at least one property which is different when compared to the same property of the precursor subtilisin.

5 These modified properties fall into several categories including: oxidative stability, substrate specificity, thermal stability, alkaline stability, catalytic activity, pH activity profile, resistance to proteolytic degradation, Km, kcat and Km/kcat ratio.

Carbonyl hydrolases are enzymes which hydrolyze compounds containing

10



15

bonds in which X is oxygen or nitrogen. They include naturally-occurring carbonyl hydrolases and recombinant carbonyl hydrolases. Naturally occurring carbonyl hydrolases principally include hydrolases, e.g. lipases and peptide hydrolases, e.g. subtilisins or metalloproteases. Peptide hydrolases include α -aminoacylpeptide hydrolase, peptidylamino-acid hydrolase, acylamino hydrolase, serine carboxypeptidase, metallocarboxypeptidase, thiol proteinase, carboxylproteinase and metalloproteinase. Serine, metallo, thiol and acid proteases are included, as well as endo and exoproteases.

"Recombinant carbonyl hydrolase" refers to a carbonyl hydrolase in which the DNA sequence encoding the naturally occurring carbonyl hydrolase is modified to produce a mutant DNA sequence which encodes the substitution, insertion or deletion of one or more amino acids in the carbonyl hydrolase amino acid sequence. Suitable modification methods are disclosed herein and in EPO Publication No. 0130756 published January 9, 1985.

Subtilisins are bacterial carbonyl hydrolases which generally act to cleave peptide bonds of proteins or peptides. As used herein, "subtilisin" means a naturally occurring subtilisin or a recombinant subtilisin. A series of naturally occurring subtilisins is known to be produced and often secreted by various bacterial species. Amino acid sequences of the members of this series are not entirely homologous. However, the subtilisins in this series exhibit the same or similar type of proteolytic activity. This class of serine proteases shares a common amino acid sequence defining a catalytic triad which distinguishes them from the chymotrypsin related class of serine proteases. The subtilisins and chymotrypsin related serine proteases both have a catalytic triad comprising aspartate, histidine and serine. In the subtilisin related proteases the relative order of these amino acids, reading from the amino to carboxy terminus is aspartate-histidineserine. In the chymotrypsin related proteases the relative order, however is histidine-aspartate-serine. Thus, subtilisin herein refers to a serine protease having the catalytic triad of subtilisin related proteases.

"Recombinant subtilisin" refers to a subtilisin in which the DNA sequence encoding the subtilisin is modified to produce a mutant DNA sequence which encodes the substitution, deletion or insertion of one or more amino acids in the naturally occurring subtilisin amino acid sequence. Suitable methods to produce such modification include those disclosed herein and in EPO Publication No. 0130756. For example, the subtilisin multiple mutant herein containing the substitution of methionine at amino acid residues 50, 124 and 222 with phenylalanine, isoleucine and glutamine, respectively, can be considered to be derived from the recombinant subtilisin containing the substitution of glutamine at residue 222 (Q222) disclosed in EPO Publication No. 0130756. The multiple mutant thus is produced by the substitution of phenylalanine for methionine at residue 50 and isoleucine for methionine at residue 124 in the Q222 recombinant subtilisin.

"Carbonyl hydrolases" and their genes may be obtained from many procaryotic and eucaryotic organisms. Suitable examples of procaryotic organisms include gram negative organisms such as *E. coli* or *pseudomonas* and gram positive bacteria such as *micrococcus* or *bacillus*. Examples of eucaryotic organisms from which carbonyl hydrolase and their genes may be obtained include yeast such as *S. cerevisiae*, fungi such as *Aspergillus* sp., and non-human mammalian sources such as, for example, Bovine sp. from which the gene encoding the carbonyl hydrolase chymosin can be obtained. As with subtilisins, a series of carbonyl hydrolases can be obtained from various related species which have amino acid sequences which are not entirely homologous between the members of that series but which nevertheless exhibit the same or similar type of biological activity. Thus, non-human carbonyl hydrolase as used herein has a functional definition which refers to carbonyl hydrolases which are associated, directly or indirectly, with procaryotic and non-human eucaryotic sources.

A "carbonyl hydrolase mutant" has an amino acid sequence which is derived from the amino acid sequence of a non-human "precursor carbonyl hydrolase". The precursor carbonyl hydrolases include naturally-occurring carbonyl hydrolases and recombinant carbonyl hydrolases. The amino acid sequence of the carbonyl hydrolase mutant is "derived" from the precursor hydrolase amino acid sequence by the substitution, deletion or insertion of one or more amino acids of the precursor amino acid sequence. Such modification is of the "precursor DNA sequence" which encodes the amino acid sequence of the precursor carbonyl hydrolase rather than manipulation of the precursor carbonyl hydrolase *per se*. Suitable methods for such manipulation of the precursor DNA sequence include methods disclosed herein and in EPO Publication No. 0130756.

Specific residues of B. amyloliquefaciens subtilisin are identified for substitution, insertion or deletion. These amino acid position numbers refer to those assigned to the B. amyloliquefaciens subtilisin sequence presented in Fig. 1. The invention, however, is not limited to the mutation of this particular subtilisin but extends to precursor carbonyl hydrolases containing amino acid residues which are "equivalent" to the particular identified residues in B. amyloliquefaciens subtilisin.

A residue (amino acid) of a precursor carbonyl hydrolase is equivalent to a residue of B. amyloliquefaciens subtilisin if it is either homologous (i.e., corresponding in position in either primary or tertiary structure) or analagous to a specific residue or portion of that residue in B. amyloliquefaciens subtilisin (i.e., having the same or similar functional capacity to combine, react, or interact chemically).

In order to establish homology to primary structure, the amino acid sequence of a precursor carbonyl hydrolase is directly compared to the B. amyloliquefaciens subtilisin primary sequence and particularly to a set of residues known to be invariant in all subtilisins for which sequence is known (Figure 5C). After aligning the conserved residues, allowing for necessary insertions and deletions in order to maintain alignment (i.e., avoiding the elimination of conserved residues through arbitrary deletion and insertion), the residues equivalent to particular amino acids in the primary sequence of B. amyloliquefaciens subtilisin are defined. Alignment of conserved residues preferably should conserve 100% of such residues. However, alignment of greater than 75% or as little as 50% of conserved residues is also adequate to define equivalent residues. Conservation of the catalytic triad, Asp32/His64/Ser221 should be maintained.

For example, in Figure 5A the amino acid sequence of subtilisin from B. amyloliquefaciens B. subtilis var. 1168 and B. licheniformis (carlsbergensis) are aligned to provide the maximum amount of homology between amino acid sequences. A comparison of these sequences shows that there are a number of conserved residues contained in each sequence. These residues are identified in Fig. 5C.

These conserved residues thus may be used to define the corresponding equivalent amino acid residues of B. amyloliquefaciens subtilisin in other carbonyl hydrolases such as thermitase derived from Thermoactinomyces. These two particular sequences are aligned in Fig. 5B to produce the maximum homology of conserved residues. As can be seen there are a number of insertions and deletions in the thermitase sequence as compared to B. amyloliquefaciens subtilisin. Thus, in thermitase the equivalent amino acid of Tyr217 in B. amyloliquefaciens subtilisin is the particular lysine shown beneath Tyr217.

In Fig. 5A, the equivalent amino acid at position 217 in B. amyloliquefaciens subtilisin is Tyr. Likewise, in B. subtilis subtilisin position 217 is also occupied by Tyr but in B. licheniformis position 217 is occupied by Leu.

Thus, these particular residues in thermitase, and subtilisin from B. subtilis and B. licheniformis may be substituted by a different amino acid to produce a mutant carbonyl hydrolase since they are equivalent in primary structure to Tyr217 in B. amyloliquefaciens subtilisin. Equivalent amino acids of course are not limited to those for Tyr217 but extend to any residue which is equivalent to a residue in B. amyloliquefaciens whether such residues are conserved or not.

Equivalent residues homologous at the level of tertiary structure for a precursor carbonyl hydrolase whose tertiary structure has been determined by x-ray crystallography, are defined as those for which the atomic coordinates of 2 or more of the main chain atoms of a particular amino acid residue of the precursor carbonyl hydrolase and B. amyloliquefaciens subtilisin (N on N, CA on CA, C on C, and O on O) are within 0.13nm and preferably 0.1nm after alignment. Alignment is achieved after the best model has been oriented and positioned to give the maximum overlap of atomic coordinates of non-hydrogen protein atoms of the carbonyl hydrolase in question to the B. amyloliquefaciens subtilisin. The best model is the crystallographic model giving the lowest R factor for experimental diffraction data at the highest resolution available.

$$R \text{ factor} = \frac{\sum_h |F_o(h)| - |F_c(h)|}{\sum_h |F_o(h)|}$$

5

Equivalent residues which are functionally analogous to a specific residue of B. amyloliquefaciens subtilisin are defined as those amino acids of the precursor carbonyl hydrolases which may adopt a conformation such that they either alter, modify or contribute to protein structure, substrate binding or catalysis in a manner defined and attributed to a specific residue of the B. amyloliquefaciens subtilisin as described herein. Further, they are those residues of the precursor carbonyl hydrolase (for which a tertiary structure has been obtained by x-ray crystallography), which occupy an analogous position to the extent that although the main chain atoms of the given residue may not satisfy the criteria of equivalence on the basis of occupying a homologous position, the atomic coordinates of at least two of the side chain atoms of the residue lie within 0.13nm of the corresponding side chain atoms of B. amyloliquefaciens subtilisin. The three dimensional structures would be aligned as outlined above.

Some of the residues identified for substitution, insertion or deletion are conserved residues whereas others are not. In the case of residues which are not conserved, the replacement of one or more amino acids is limited to substitutions which produce a mutant which has an amino acid sequence that does not correspond to one found in nature. In the case of conserved residues, such replacements should not result in a naturally occurring sequence. The carbonyl hydrolase mutants of the present invention include the mature forms of carbonyl hydrolase mutants as well as the pro- and prepro-forms of such hydrolase mutants. The prepro-forms are the preferred construction since this facilitates the expression, secretion and maturation of the carbonyl hydrolase mutants.

"Expression vector" refers to a DNA construct containing a DNA sequence which is operably linked to a suitable control sequence capable of effecting the expression of said DNA in a suitable host. Such control sequences include a promoter to effect transcription, an optional operator sequence to control such transcription, a sequence encoding suitable mRNA ribosome binding sites, and sequences which control termination of transcription and translation. The vector may be a plasmid, a phage particle, or simply a potential genomic insert. Once transformed into a suitable host, the vector may replicate and function independently of the host genome, or may, in some instances, integrate into the genome itself. In the present specification, "plasmid" and "vector" are sometimes used interchangeably as the plasmid is the most commonly used form of vector at present. However, the invention is intended to include such other forms of expression vectors which serve equivalent functions and which are, or become, known in the art.

The "host cells" used in the present invention generally are procaryotic or eucaryotic hosts which preferably have been manipulated by the methods disclosed in EPO Publication No. 0130756 to render them incapable of secreting enzymatically active endoprotease. A preferred host cell for expressing subtilisin is the *Bacillus* strain BG2036 which is deficient in enzymatically active neutral protease and alkaline protease (subtilisin). The construction of strain BG2036 is described in detail in EPO Publication No. 0130756 and further described by Yang, M.Y., et al. (1984) *J. Bacteriol.* 160, 15-21. Other host cells for expressing subtilisin include *Bacillus subtilis* I168 (EPO Publication No. 0130756).

Host cells are transformed or transfected with vectors constructed using recombinant DNA techniques. Such transformed host cells are capable of either replicating vectors encoding the carbonyl hydrolase mutants or expressing the desired carbonyl hydrolase mutant. In the case of vectors which encode the pre or prepro form of the carbonyl hydrolase mutant, such mutants, when expressed, are typically secreted from the host cell into the host cell medium.

"Operably linked" when describing the relationship between two DNA regions simply means that they are functionally related to each other. For example, a presequence is operably linked to a peptide if it functions as a signal sequence, participating in the secretion of the mature form of the protein most probably involving cleavage of the signal sequence. A promoter is operably linked to a coding sequence if it controls the transcription of the sequence; a ribosome binding site is operably linked to a coding sequence if it is positioned so as to permit translation.

The genes encoding the naturally-occurring precursor carbonyl hydrolase may be obtained in accord with the general methods described herein in EPO publication No. 0130756.

Once the carbonyl hydrolase gene has been cloned, a number of modifications are undertaken to enhance the use of the gene beyond synthesis of the naturally-occurring precursor carbonyl hydrolase. Such modifications include the production of recombinant carbonyl hydrolases as disclosed in EPO

Publication No. 0130756 and the production of carbonyl hydrolase mutants described herein.

The carbonyl hydrolase mutants of the present invention may be generated by site specific mutagenesis (Smith, M. (1985) *Ann. Rev. Genet.* 423; Zoeller, M.J., et al. (1982) *Nucleic Acid Res.* 10, 6487-6500), cassette mutagenesis (EPO Publication No. 0130756) or random mutagenesis (Shortle, D., et al. (1985) *Genetics*, 110, 539; Shortle, D., et al. (1986) *Proteins: Structure, Function and Genetics*, 1, 81; Shortle, D. (1986) *J. Cell. Biochem.*, 30, 281; Alber, T., et al. (1985) *Proc. Natl. Acad. of Sci.*, 82, 747; Matsumura, M., et al. (1985) *J. Biochem.*, 260, 15298; Liao, H., et al. (1986) *Proc. Natl. Acad. of Sci.*, 83, 576) of the cloned precursor carbonyl hydrolase. Cassette mutagenesis and the random mutagenesis method disclosed herein are preferred.

The mutant carbonyl hydrolases expressed upon transformation of suitable hosts are screened for enzymes exhibiting one or more properties which are substantially different from the properties of the precursor carbonyl hydrolases, e.g., changes in substrate specificity, oxidative stability, thermal stability, alkaline stability, resistance to proteolytic degradation, pH-activity profiles and the like.

A change in substrate specificity is defined as a difference between the k_{cat}/K_m ratio of the precursor carbonyl hydrolase and that of the hydrolase mutant. The k_{cat}/K_m ratio is a measure of catalytic efficiency. Carbonyl hydrolase mutants with increased or diminished k_{cat}/K_m ratios are described in the examples. Generally, the objective will be to secure a mutant having a greater (numerically large) k_{cat}/K_m ratio for a given substrate, thereby enabling the use of the enzyme to more efficiently act on a target substrate. A substantial change in k_{cat}/K_m ratio is preferably at least 2-fold increase or decrease. However, smaller increases or decreases in the ratio (e.g., at least 1.5-fold) are also considered substantial. An increase in k_{cat}/K_m ratio for one substrate may be accompanied by a reduction in k_{cat}/K_m ratio for another substrate. This is a shift in substrate specificity, and mutants exhibiting such shifts have utility where the precursor hydrolase is undesirable, e.g. to prevent undesired hydrolysis of a particular substrate in an admixture of substrates. K_m and k_{cat} are measured in accord with known procedures, as described in EPO Publication No. 0130756 or as described herein.

Oxidative stability is measured either by known procedures or by the methods described hereinafter. A substantial change in oxidative stability is evidenced by at least about 50% increase or decrease (preferably decrease) in the rate of loss of enzyme activity when exposed to various oxidizing conditions. Such oxidizing conditions are exposure to the organic oxidant diperoxidodecanoic acid (DPDA) under the conditions described in the examples.

Alkaline stability is measured either by known procedures or by the methods described herein. A substantial change in alkaline stability is evidenced by at least about a 5% or greater increase or decrease (preferably increase) in the half life of the enzymatic activity of a mutant when compared to the precursor carbonyl hydrolase. In the case of subtilisins, alkaline stability was measured as a function of autoproteolytic degradation of subtilisin at alkaline pH, e.g. for example, 0.1M sodium phosphate, pH 12 at 25° or 30° C.

Thermal stability is measured either by known procedures or by the methods described herein. A substantial change in thermal stability is evidenced by at least about a 5% or greater increase or decrease (preferably increase) in the half-life of the catalytic activity of a mutant when exposed to a relatively high temperature and neutral pH as compared to the precursor carbonyl hydrolase. In the case of subtilisins, thermal stability is measured by the autoproteolytic degradation of subtilisin at elevated temperatures and neutral pH, e.g., for example 2mM calcium chloride, 50mM MOPS pH 7.0 at 59° C.

The inventors have produced mutant subtilisins containing the substitution of the amino acid residues of *B. amyloliquefaciens* subtilisin shown in Table I. The wild type amino acid sequence and DNA sequence of *B. amyloliquefaciens* subtilisin is shown in Fig. 1.

TABLE I

	Residue	Replacement Amino Acid
5	Tyr21	F A
	Thr22	C
	Ser24	C
	Asp32	Q S
	Ser33	A T
10	Asp36	A G
	Gly46	V
	Ala48	E V R
	Ser49	C L
	Met50	C F V
15	Asn77	D
	Ser87	C
	Lys94	C
	Val95	C
	Leu96	D
20	Tyr104	A C D E F G H I K L M N P Q R S T V W
	Ile107	V
	Gly110	C R
	Met124	I L
	Asn155	A D H Q T
25	Glu156	Q S
	Gly166	C E I L M P S T W Y
	Gly169	C D E F H I K L M N P Q R T V W Y
	Lys170	E R
	Tyr171	F
30	Pro172	E Q
	Phel89	A C D E G H I K L M N P Q R S T V W Y
	Asp197	R A
	Met199	I
	Ser204	C R L P
35	Lys213	R T
	Tyr217	A C D E F G H I K L M N P Q R S T V W
	Ser221	A C

40 The different amino acids substituted are represented in Table I by the following single letter designations:

45

50

55

Amino acid or residue thereof	3-letter symbol	1-letter symbol
Alanine	Ala	A
Glutamate	Glu	E
Glutamine	Gln	Q
Aspartate	Asp	D
Asparagine	Asn	N
Leucine	Leu	L
Glycine	Gly	G
Lysine	Lys	K
Serine	Ser	S
Valine	Val	V
Arginine	Arg	R
Threonine	Thr	T
Proline	Pro	P
Isoleucine	Ile	I
Methionine	Met	M
Phenylalanine	Phe	F
Tyrosine	Tyr	Y
Cysteine	Cys	C
Tryptophan	Trp	W
Histidine	His	H

Except where otherwise indicated by context, wild-type amino acids are represented by the above three-letter symbols and replaced amino acids by the above single-letter symbols. Thus, if the methionine at residue 50 in *B. amyloliquefaciens* subtilisin is replaced by phenylalanine, this mutation (mutant) may be designated Met50F or F50. Similar designations are used for multiple mutants.

In addition to the amino acids used to replace the residues disclosed in Table I, other replacements of amino acids at these residues are expected to produce mutant subtilisins having useful properties. These residues and replacement amino acids are shown in Table II.

TABLE II

	Residue	Replacement Amino Acid(s)
5	Tyr-21	L
	Thr22	K
	Ser24	A
	Asp32	
	Ser33	G
10	Gly46	
	Ala48	
	Ser49	
	Met50	L K I V
	Asn77	D
15	Ser87	N
	Lys94	R Q
	Val95	L I
	Tyr104	
	Met124	K A
20	Ala152	C L I T M
	Asn155	
	Glu156	A T M L Y
	Gly166	
	Gly169	
25	Tyr171	K R E Q
	Pro172	D N
	Phe189	
	Tyr217	
	Ser221	
30	Met222	

Each of the mutant subtilisins in Table I contain the replacement of a single residue of the B. amyloliquefaciens amino acid sequence. These particular residues were chosen to probe the influence of such substitutions on various properties of B. amyloliquefacien subtilisin.

Thus, the inventors have identified Met124 and Met222 as important residues which if substituted with another amino acid produce a mutant subtilisin with enhanced oxidative stability. For Met124, Leu and Ile are preferred replacement amino acids. Preferred amino acids for replacement of Met222 are disclosed in EPO Publication No. 0130756.

Various other specific residues have also been identified as being important with regard to substrate specificity. These residues include Tyr104, Ala152, Glu156, Gly166, Gly169, Phe189 and Tyr217 for which mutants containing the various replacement amino acids presented in Table I have already been made, as well as other residues presented below for which mutants have yet to be made.

The identification of these residues, including those yet to be mutated, is based on the inventors' high resolution crystal structure of B. amyloliquefaciens subtilisin to 1.8 Å (see Table III), their experience with in vitro mutagenesis of subtilisin and the literature on subtilisin. This work and the x-ray crystal structures of subtilisin containing covalently bound peptide inhibitors (Robertus, J.D., et al. (1972) Biochemistry 11, 2439-2449), product complexes (Robertus, J.D., et al. (1972) Biochemistry 11, 4293-4303), and transition state analogs (Matthews, D.A., et al (1975) J. Biol. Chem. 250, 7120-7126; Poulos, T.L., et al. (1976) J. Biol. Chem. 251, 1097-1103), has helped in identifying an extended peptide binding cleft in subtilisin. This substrate binding cleft together with substrate is schematically diagrammed in Fig. 2, according to the nomenclature of Schechter, I., et al. (1967) Biochem Bio. Res. Commun. 27, 157. The scissile bond in the substrate is identified by an arrow. The P and P' designations refer to the amino acids which are positioned respectively toward the amino or carboxy terminus relative to the scissile bond. The S and S' designations refer to subsites in the substrate binding cleft of subtilisin which interact with the corresponding substrate amino acid residues.

Atomic Coordinates for the
Apoenzyme Form of *B. Amyloliqnefaciens*
Subtilisin to 1.8Å Resolution

5

10

15

20

25

30

35

40

45

50

55

1	ALA M	19.434	53.195	-21.754	1	ALA CA	19.811	51.774	-21.965
1	ALA C	18.731	50.925	-21.324	1	ALA O	18.374	51.197	-20.175
1	ALA CB	23.099	51.518	-21.183	2	GLN M	18.268	49.884	-22.041
2	GLN CA	17.219	49.008	-21.434	2	GLN C	17.875	47.706	-20.992
2	GLN O	18.765	47.165	-21.691	2	GLN CB	16.125	48.760	-22.449
2	GLN CG	15.328	47.805	-21.927	2	GLN CD	13.912	47.762	-22.930
2	GLN OE1	13.023	48.612	-22.067	2	GLN ME2	14.115	46.917	-23.924
3	SER M	17.477	47.205	-19.852	3	SER CA	17.950	45.868	-19.437
3	SER C	16.735	44.918	-19.490	3	SER O	15.590	45.352	-19.729
3	SER CB	18.588	45.838	-18.069	3	SER OG	17.482	44.210	-17.049
4	VAL M	16.991	43.646	-19.725	4	VAL CA	15.946	42.619	-19.639
4	VAL C	16.129	41.934	-18.298	4	VAL O	17.123	41.178	-18.086
4	VAL CB	16.008	41.622	-20.822	4	VAL CG1	14.874	40.572	-20.741
4	VAL CG2	16.037	42.266	-22.186	5	PRO M	15.239	42.104	-17.331
5	PRO CA	15.384	41.415	-16.827	5	PRO C	15.501	39.905	-16.249
5	PRO O	14.885	39.263	-17.144	5	PRO CS	14.150	41.880	-15.263
5	PRO CG	13.841	43.215	-15.921	5	PRO CD	14.044	42.986	-17.417
6	TYR M	16.363	39.248	-15.487	6	TYR CA	16.428	37.883	-15.715
6	TYR C	15.359	36.975	-15.528	6	TYR O	15.224	35.943	-16.235
6	TYR CB	17.824	37.323	-14.834	6	TYR CG	18.021	35.847	-15.055
6	TYR CD1	18.437	35.452	-16.346	6	TYR CD2	17.696	34.988	-14.071
6	TYR CE1	18.535	34.870	-16.653	6	TYR CE2	17.815	33.539	-14.379
6	TYR C2	18.222	33.154	-15.628	6	TYR OH	18.312	31.838	-15.996
7	GLY M	14.464	37.362	-14.630	7	GLY CA	13.211	36.640	-14.376
7	GLY C	12.480	36.535	-15.670	7	GLY O	11.747	35.478	-15.883
8	VAL M	12.441	37.529	-16.541	8	VAL CA	11.777	37.523	-17.836
8	VAL C	12.363	36.433	-18.735	8	VAL O	11.639	35.716	-19.470
8	VAL CB	11.765	38.900	-18.567	8	VAL CG1	12.186	38.893	-19.943
8	VAL CG2	18.991	39.919	-17.733	9	SER M	13.641	36.318	-18.775
9	SER CA	14.419	35.342	-19.562	9	SER C	14.188	33.920	-18.945
9	SER O	14.112	33.014	-19.801	9	SER CB	15.926	35.632	-19.585
9	SER OG	16.162	36.747	-20.358	10	GLN M	14.115	33.887	-17.662
10	GLN CA	13.964	32.636	-16.876	10	GLN C	12.687	31.887	-17.277
10	GLN O	12.785	30.647	-17.413	10	GLN CB	14.125	32.885	-15.410
10	GLN CG	16.295	31.617	-14.588	10	GLN CD	14.486	31.911	-13.147
10	GLN OE1	14.554	33.868	-12.744	10	GLN ME2	14.552	30.960	-12.251
11	ILE M	11.625	32.575	-17.670	11	ILE CA	10.373	31.984	-18.182
11	ILE C	10.209	31.792	-19.605	11	ILE O	9.171	31.333	-20.180
11	ILE CB	9.132	32.669	-17.475	11	ILE CG1	9.846	34.117	-18.849
11	ILE CG2	9.162	32.655	-15.941	11	ILE CD1	7.588	34.648	-17.923
12	LYS M	11.272	32.185	-20.277	12	LYS CA	11.388	32.119	-21.722
12	LYS C	10.456	33.886	-22.522	12	LYS O	10.178	32.783	-23.686
12	LYS CB	11.257	30.646	-22.216	12	LYS CG	12.283	29.830	-21.423
12	LYS CD	12.543	28.517	-22.159	12	LYS CE	13.023	27.467	-21.166
12	LYS ME2	14.476	27.680	-20.935	13	ALA M	10.189	34.138	-21.991
13	ALA CA	9.325	35.198	-22.631	13	ALA C	10.026	35.716	-23.863
13	ALA O	9.338	35.804	-24.981	13	ALA CB	8.885	36.195	-21.565
14	PRO M	11.332	35.950	-23.893	14	PRO CA	11.985	36.430	-23.128
14	PRO C	11.786	35.957	-26.317	14	PRO O	11.778	36.847	-27.645
14	PRO CB	13.462	36.580	-24.692	14	PRO CG	13.328	36.978	-23.221
14	PRO CD	12.281	35.936	-22.758	15	ALA M	11.560	34.236	-24.179
15	ALA CA	11.379	33.458	-27.367	15	ALA C	10.082	33.795	-28.052
15	ALA O	10.808	33.718	-29.278	15	ALA CB	11.552	31.969	-27.862
16	LEU M	9.885	34.138	-27.240	16	LEU CA	7.791	34.558	-27.828
16	LEU C	7.912	35.925	-28.521	16	LEU O	7.342	36.126	-29.588
16	LEU CB	6.746	34.673	-26.698	16	LEU CG	9.798	33.665	-26.522
16	LEU CD1	9.881	33.234	-27.889	16	LEU CD2	6.694	32.287	-26.283
17	THR M	8.665	36.828	-27.922	17	THR CA	8.898	38.151	-28.938
17	THR C	9.510	37.981	-29.890	17	THR O	9.187	38.622	-30.856
17	THR CB	9.788	39.188	-27.652	17	THR CG	9.185	39.288	-26.262
17	THR CD1	9.930	39.887	-25.272	17	THR CD2	8.006	38.926	-25.894
17	THR CE1	9.226	39.914	-24.144	17	THR ME2	8.879	39.328	-24.381
18	SER M	10.443	37.833	-30.822	18	SER CA	11.189	36.739	-31.322

EP 0 251 446 B1

18	SEP C	19.119	36.123	-32.353	19	SEP D	19.849	36.112	-33.834
18	SEP CD	12.311	35.799	-31.172	19	SEP DC	13.121	36.488	-39.199
19	GLN M	9.880	35.685	-31.943	19	GLN CA	8.982	36.962	-32.878
19	GLN C	7.142	36.111	-33.303	19	GLN D	6.297	36.972	-36.219
19	GLN CB	7.221	33.849	-32.280	19	GLN CC	7.979	32.882	-31.823
19	GLN CD	6.823	31.707	-31.181	19	GLN DC1	9.719	31.833	-31.444
19	GLN ME2	7.362	38.832	-30.256	20	GLY M	7.285	37.223	-32.887
20	GLY CA	6.369	38.387	-32.859	20	GLY C	5.181	38.492	-31.888
20	GLY D	4.263	39.276	-32.215	21	TYR M	5.282	37.801	-38.761
21	TYR CA	4.118	37.831	-29.763	21	TYR C	4.879	38.532	-28.923
21	TYR D	5.422	38.974	-27.756	21	TYR CB	3.488	36.431	-29.463
21	TYR CG	2.973	31.784	-30.788	21	TYR CD1	1.795	36.332	-31.238
21	TYR CD2	3.650	36.794	-31.397	21	TYR CE1	1.306	33.797	-32.444
21	TYR CE2	3.193	34.261	-32.888	21	TYR CE	2.883	34.753	-33.867
21	TYR CM	1.301	36.241	-34.750	22	TYR M	3.902	39.680	-38.284
22	TYR CA	4.262	40.927	-27.129	22	TYR C	3.891	40.922	-26.244
22	TYR D	3.287	41.725	-25.325	22	TYR CB	3.133	41.759	-27.611
22	TYR DC1	4.319	42.457	-25.197	22	TYR CC2	4.476	41.323	-28.229
22	GLY M	1.839	48.285	-26.453	23	GLY CA	8.899	48.600	-21.342
23	GLY C	-0.137	41.631	-26.118	23	GLY D	-1.013	42.895	-25.319
24	SEP M	-8.823	41.967	-27.371	24	SEP CA	-8.897	42.987	-28.812
24	SEP C	-2.383	42.626	-27.864	24	SEP D	-2.813	41.508	-28.168
24	SEP CB	-8.734	43.120	-29.320	24	SEP DC	8.563	43.632	-29.728
25	ASN M	-3.859	43.692	-27.515	25	ASN CA	-4.519	43.687	-27.393
25	ASN C	-3.813	42.873	-26.203	25	ASN D	-6.233	42.648	-26.199
25	ASN CB	-3.163	43.227	-28.705	25	ASN CC	-4.960	44.178	-29.883
25	ASN CD1	-4.963	43.767	-31.083	25	ASN CD2	-4.747	43.461	-29.594
26	VAL M	-4.177	42.649	-25.292	26	VAL CA	-4.674	41.679	-24.143
26	VAL C	-4.792	42.832	-22.937	26	VAL D	-3.858	43.419	-27.689
26	VAL CB	-3.714	40.903	-23.821	26	VAL CC1	-4.160	39.802	-22.548
26	VAL CC2	-3.998	39.576	-25.818	27	LYS M	-3.910	42.613	-22.301
27	LYS CA	-6.133	43.824	-21.175	27	LYS C	-5.818	42.872	-18.841
27	LYS D	-6.403	41.873	-19.413	27	LYS CB	-7.990	43.981	-21.149
27	LYS CG	-8.046	44.573	-22.490	27	LYS CD	-9.321	45.302	-22.820
27	LYS CE	-10.304	45.497	-23.137	27	LYS CE1	-9.686	46.253	-24.264
28	VAL M	-4.818	43.462	-19.205	28	VAL CA	-4.497	42.938	-17.897
28	VAL C	-4.753	43.959	-18.828	28	VAL D	-4.289	45.893	-16.817
28	VAL CB	-2.926	42.666	-17.932	28	VAL CC1	-2.466	42.183	-16.589
28	VAL CC2	-2.667	41.885	-19.173	29	ALA M	-5.484	43.527	-15.813
29	ALA CA	-5.747	44.330	-14.639	29	ALA C	-6.790	44.818	-13.513
29	ALA D	-4.666	42.843	-13.104	29	ALA CB	-7.172	46.187	-14.181
30	VAL M	-4.857	45.033	-13.872	30	VAL CA	-3.146	44.962	-11.918
30	VAL C	-3.938	45.489	-10.681	30	VAL D	-4.185	46.648	-18.978
30	VAL CB	-1.886	45.810	-12.149	30	VAL CC1	-9.996	45.981	-18.988
30	VAL CC2	-1.833	45.236	-13.307	31	ILE M	-4.914	44.515	-9.877
31	ILE CA	-5.328	46.846	-8.679	31	ILE C	-4.346	44.933	-7.546
31	ILE D	-3.824	43.913	-6.987	31	ILE CB	-6.457	43.774	-8.581
31	ILE CC1	-7.293	43.787	-9.799	31	ILE CC2	-7.278	46.838	-7.223
31	ILE CD1	-8.617	42.836	-9.717	32	ASP M	-6.844	46.183	-7.217
32	ASP CA	-2.944	46.467	-6.255	32	ASP C	-3.871	47.889	-5.765
32	ASP D	-4.197	48.418	-8.382	32	ASP CB	-1.493	46.129	-7.892
32	ASP CG	-8.483	45.782	-6.273	32	ASP CD1	8.834	44.592	-6.576
32	ASP CD2	-8.881	46.419	-5.330	33	SER M	-1.931	48.512	-3.394
33	SER CA	-1.893	49.837	-4.881	33	SER C	-1.982	88.976	-3.888
33	SER D	-1.786	82.134	-5.363	33	SER CB	-8.621	49.922	-3.839
33	SER DC	8.333	80.825	-4.774	34	GLY M	-2.173	58.740	-7.884
34	GLY CA	-2.233	51.728	-8.163	34	GLY C	-1.935	51.648	-9.857
34	GLY D	-8.144	50.831	-8.761	35	ILE M	-8.963	52.431	-18.182
35	ILE CA	8.388	52.438	-18.993	35	ILE C	8.368	53.919	-11.263
35	ILE D	-8.327	54.638	-21.744	35	ILE CB	-8.842	51.694	-12.367
35	ILE CC1	-8.538	50.210	-12.897	35	ILE CC2	3.149	51.741	-13.382
35	ILE CD1	-8.962	49.483	-13.424	36	ASP M	1.816	54.253	-18.971
36	ASP CA	2.359	55.618	-11.232	36	ASP C	2.281	55.956	-12.782

45

50

55

5	36	ASP D	3.004	55.471	-13.579	36	ASP CB	3.712	55.720	-10.516
	36	ASP CG	4.339	57.099	-10.884	36	ASP DD1	3.755	57.974	-11.429
	36	ASP DD2	5.440	57.277	-10.263	37	SER M	1.304	56.022	-13.111
	37	SER CA	1.103	57.221	-14.512	37	SER C	2.377	56.093	-14.949
	37	SER O	2.543	58.303	-16.151	37	SER CB	-0.093	58.049	-14.788
	37	SER DG	-0.090	59.133	-13.870	38	SER M	3.163	58.614	-14.001
	38	SER CA	4.261	59.505	-14.487	38	SER C	5.466	58.705	-14.992
	38	SER O	6.543	59.251	-15.285	38	SER CB	4.742	60.435	-13.398
	38	SER DG	5.376	59.865	-12.234	39	MIS M	5.454	57.390	-14.892
	39	MIS CA	6.637	56.574	-15.291	39	MIS C	6.681	56.481	-16.778
	39	MIS O	5.730	55.078	-17.419	39	MIS CB	6.637	55.203	-14.515
	39	MIS CG	0.014	56.609	-14.456	39	MIS DD1	0.795	54.356	-13.561
	39	MIS DD2	0.769	56.345	-13.389	39	MIS DD1	9.970	53.930	-15.130
10	39	MIS DD2	9.986	53.910	-13.908	40	PRO M	7.007	56.836	-17.387
	40	PRO CA	7.988	56.697	-18.031	40	PRO C	0.156	55.280	-19.357
	40	PRO O	0.032	55.097	-20.570	40	PRO CB	9.247	57.533	-19.161
	40	PRO CG	10.053	57.485	-17.902	40	PRO CD	0.988	57.452	-16.776
	41	ASP M	0.481	54.328	-18.485	41	ASP DD2	11.148	58.399	-18.668
	41	ASP DD1	10.325	51.395	-20.429	41	ASP CG	10.473	51.307	-19.211
	41	ASP CB	9.799	52.239	-18.224	41	ASP CA	0.645	52.959	-18.966
	41	ASP C	7.311	52.163	-18.839	41	ASP O	7.396	50.947	-18.977
15	42	LEU M	6.185	52.003	-18.558	42	LEU CA	4.892	52.147	-18.466
	42	LEU C	3.924	52.907	-19.376	42	LEU O	3.993	54.163	-19.490
	42	LEU CB	4.421	52.158	-17.008	42	LEU CG	5.102	51.363	-15.946
	42	LEU CD1	4.535	51.546	-14.581	42	LEU CD2	5.273	49.877	-16.358
	43	LVS M	3.018	52.135	-19.946	43	LVS CA	2.093	52.685	-20.721
	43	LVS C	0.637	52.156	-20.818	43	LVS O	0.504	50.920	-19.820
	43	LVS CB	2.021	52.389	-22.169	43	LVS CG	0.685	52.436	-22.910
	43	LVS CD	0.998	52.062	-24.339	43	LVS CE	-0.180	52.504	-25.260
20	43	LVS M2	0.337	51.757	-26.418	44	VAL M	-0.191	53.835	-19.490
	44	VAL CA	-1.487	52.639	-18.765	44	VAL C	-2.571	52.087	-19.731
	44	VAL O	-2.623	53.986	-20.434	44	VAL CB	-1.480	53.351	-17.383
	44	VAL CG1	-2.724	52.941	-16.582	44	VAL CG2	-0.197	53.194	-16.553
	45	ALA M	-3.494	51.951	-19.071	45	ALA CA	-4.619	51.977	-20.810
	45	ALA C	-5.841	52.507	-20.053	45	ALA O	-6.703	53.085	-20.783
	45	ALA CB	-4.831	50.580	-21.389	46	GLY M	-5.910	52.354	-18.748
	46	GLY CA	-7.082	52.837	-18.081	46	GLY C	-6.987	52.443	-18.538
25	46	GLY O	-5.938	52.006	-16.035	47	GLY M	-8.092	52.658	-15.793
	47	GLY CA	-8.014	52.246	-14.388	47	GLY C	-9.179	52.757	-13.572
	47	GLY O	-9.988	53.481	-14.185	48	ALA M	-9.221	52.446	-12.330
	48	ALA CA	-10.255	52.878	-11.382	48	ALA C	-9.790	52.675	-9.968
	48	ALA O	-9.046	51.720	-9.725	48	ALA CB	-11.558	52.100	-11.617
	49	SER M	-10.149	53.547	-9.837	49	SER CA	-9.752	53.355	-7.652
	49	SER C	-10.947	52.986	-6.783	49	SER O	-11.972	53.677	-6.908
	49	SER CB	-9.092	54.588	-7.029	49	SER DG	-0.879	54.255	-5.650
30	50	MET M	-10.835	52.007	-5.932	50	MET CA	-11.852	51.549	-4.974
	50	MET C	-11.463	51.962	-3.561	50	MET O	-11.997	51.398	-2.575
	50	MET CB	-12.012	50.018	-4.996	50	MET CG	-11.912	49.463	-6.389
	50	MET SD	-13.468	49.889	-7.256	50	MET CE	-12.808	50.111	-0.983
	51	VAL M	-10.427	52.760	-3.422	51	VAL CA	-9.968	53.170	-2.067
	51	VAL C	-10.630	54.562	-1.987	51	VAL O	-10.237	55.437	-2.682
	51	VAL CB	-8.443	53.155	-2.000	51	VAL CG1	-7.892	53.579	-0.631
35	51	VAL CG2	-7.764	51.815	-2.302	52	PRO M	-11.621	54.693	-1.056
	52	PRO CA	-12.372	55.933	-0.821	52	PRO C	-11.490	57.123	-0.440
	52	PRO O	-11.771	58.220	-0.925	52	PRO CB	-13.680	55.594	0.244
	52	PRO CG	-13.583	54.183	0.085	52	PRO CO	-12.264	53.620	-0.175
	53	SER M	-10.442	56.986	0.299	53	SER CA	-9.538	57.982	0.682
	53	SER C	-0.428	58.245	-0.326	53	SER O	-7.679	59.224	-0.030
	53	SER CB	-9.084	57.787	2.069	53	SER DG	-8.256	56.521	2.127
	54	GLU M	-8.254	57.623	-1.393	54	GLU CA	-7.204	57.648	-2.421
40	54	GLU C	-7.767	57.303	-3.785	54	GLU O	-7.533	56.243	-4.379
	54	GLU CB	-6.134	56.599	-2.154	54	GLU CG	-5.289	56.959	-0.927
	54	GLU CD	-6.044	54.849	-0.070	54	GLU DD1	-3.646	55.606	-1.068

45

50

55

54	GLW DE1	-3.988	55.777	0.271	55	TMR M	-0.571	58.251	-4.249
55	TMR CA	-9.433	58.121	-5.441	55	TMR C	-8.764	58.139	-4.779
55	TMR D	-9.433	57.919	-7.818	55	TMR CB	-18.386	59.208	-5.383
55	TMR CG1	-9.885	60.510	-5.418	55	TMR CG2	-11.432	59.163	-4.017
56	ASN M	-7.482	58.483	-6.877	56	ASN MD2	-4.938	61.179	-9.881
56	ASN OD1	-5.875	58.967	-18.337	56	ASN CG	-5.273	59.925	-9.555
56	ASN CB	-5.898	59.694	-8.288	56	ASN CA	-6.762	58.425	-8.288
56	ASN C	-6.812	57.894	-8.305	56	ASN D	-5.184	56.866	-7.478
57	PRO M	-6.362	56.261	-9.258	57	PRO CG	-7.123	55.257	-11.177
57	PRO CD	-7.384	56.433	-10.272	57	PRO CB	-6.644	54.178	-10.235
57	PRO CA	-5.679	56.961	-9.332	57	PRO C	-4.381	55.882	-9.966
57	PRO D	-3.589	56.128	-9.945	58	PHE M	-3.998	56.262	-18.491
58	PHE CA	-2.747	56.577	-11.222	58	PHE C	-1.712	57.129	-10.253
58	PHE D	-0.635	57.497	-10.680	58	PHE CG	-2.943	57.582	-12.423
58	PHE CG	-3.983	56.968	-13.357	58	PHE CD1	-3.756	55.788	-14.859
58	PHE CD2	-5.211	57.630	-13.459	58	PHE CE1	-4.722	55.255	-14.928
58	PHE CE2	-6.194	57.895	-14.276	58	PHE C2	-5.949	55.939	-15.051
59	GLM M	-2.844	57.119	-8.998	59	GLM CA	-1.172	57.583	-7.934
59	GLM C	-0.887	56.483	-7.808	59	GLM D	-1.439	56.883	-6.115
59	GLM CB	-1.862	58.668	-7.889	59	GLM CG	-0.942	59.261	-6.834
59	GLM CD	-1.798	60.157	-5.150	59	GLM DE1	-1.484	61.288	-6.836
59	GLM ME2	-2.959	59.685	-6.742	60	ASP M	0.418	55.895	-7.211
60	ASP CA	0.851	54.792	-6.304	60	ASP C	1.631	55.267	-5.098
60	ASP D	2.827	55.550	-5.231	60	ASP CB	1.594	53.744	-7.188
60	ASP CG	2.877	52.538	-6.380	60	ASP OD1	1.746	52.337	-5.190
60	ASP OD2	2.915	51.841	-7.030	61	ASN M	0.959	55.265	-3.950
61	ASN MD2	-1.364	57.747	-2.347	61	ASN BD1	0.666	58.566	-2.875
61	ASN CG	-8.848	57.670	-2.399	61	ASN CB	0.531	56.481	-1.784
61	ASN CA	1.557	55.734	-2.788	61	ASN C	2.291	54.632	-1.948
61	ASN D	2.933	54.862	-8.902	62	ASN M	2.218	53.434	-2.468
62	ASN CA	2.877	52.348	-1.789	62	ASN C	4.124	51.893	-2.479
62	ASN D	4.951	51.313	-1.778	62	ASN CB	1.783	51.319	-1.421
62	ASN CG	2.371	50.183	-0.697	62	ASN OD1	2.633	49.877	-1.343
62	ASN MD2	2.622	50.288	0.601	63	SER M	4.152	52.184	-3.761
63	SER CA	5.189	51.696	-4.789	63	SER C	5.871	50.256	-5.289
63	SER D	5.593	49.790	-6.269	63	SER CB	6.523	51.958	-4.812
63	SER CG	6.871	50.698	-3.418	64	MIS M	4.282	49.475	-4.639
64	MIS CA	3.994	48.855	-4.935	64	MIS C	3.366	47.759	-4.261
64	MIS D	3.861	46.974	-7.188	64	MIS CB	3.184	47.581	-3.747
64	MIS CG	3.144	46.821	-3.726	64	MIS MD1	2.187	45.247	-4.241
64	MIS CD2	4.054	45.194	-3.135	64	MIS CE1	2.416	43.966	-4.054
64	MIS ME2	3.556	43.928	-3.368	65	GLY M	2.287	48.428	-6.587
65	GLY CA	1.552	48.264	-7.838	65	GLY C	2.392	48.636	-9.837
65	GLY D	2.238	48.878	-10.134	66	TMR M	3.233	49.659	-8.832
66	TMR CA	4.864	50.117	-9.954	66	TMR C	5.889	49.809	-10.291
66	TMR D	5.333	48.789	-11.461	66	TMR CB	4.744	51.511	-9.667
66	TMR CG1	3.637	52.425	-9.406	66	TMR CG2	5.536	52.878	-10.849
67	MIS M	5.685	48.443	-9.274	67	MIS CA	6.783	47.361	-9.458
67	MIS C	6.091	46.141	-10.143	67	MIS D	6.649	45.638	-11.158
67	MIS CB	7.388	47.871	-8.864	67	MIS CG	8.595	46.275	-8.148
67	MIS CD1	8.598	46.987	-8.276	67	MIS CD2	9.904	46.678	-8.876
67	MIS CE1	9.857	44.491	-8.299	67	MIS ME2	18.678	45.514	-8.186
68	VAL M	4.892	45.749	-9.731	68	VAL CA	4.142	44.687	-10.266
68	VAL C	3.856	44.868	-11.748	68	VAL D	4.114	43.942	-12.535
68	VAL CB	2.939	44.252	-9.386	68	VAL CG1	1.968	43.268	-18.828
68	VAL CG2	3.319	43.785	-8.888	69	ALA M	3.373	46.849	-12.113
69	ALA CA	3.837	46.468	-13.429	69	ALA C	4.193	46.398	-14.611
69	ALA D	4.828	45.913	-15.565	69	ALA CB	2.332	47.851	-13.386
70	GLY M	5.348	46.782	-13.914	70	GLY CA	6.595	48.885	-14.678
70	GLY C	7.846	45.378	-15.821	70	GLY D	7.684	43.154	-18.119
71	TMR M	6.828	44.431	-14.138	71	TMR CA	7.177	43.819	-14.446
71	TMR C	6.224	42.586	-15.563	71	TMR D	6.682	41.828	-16.895
71	TMR CB	7.119	42.878	-13.191	71	TMR CG1	8.191	42.592	-12.398

45

50

55

5	71	YMR CG2	7.274	40.983	-13.596	72	VAL M	4.930	62.007	-19.627
	72	VAL CA	3.976	42.491	-16.484	72	VAL C	4.312	63.004	-17.831
	72	VAL B	4.341	42.380	-18.868	72	VAL CB	2.916	62.867	-14.885
	72	VAL CG1	1.512	42.400	-17.170	72	VAL CG2	2.142	62.327	-14.723
	73	ALA M	4.534	44.417	-17.180	73	ALA CA	4.987	45.891	-19.167
	73	ALA C	5.433	46.333	-19.355	73	ALA B	5.062	47.188	-20.216
	73	ALA CB	3.107	45.441	-19.433	74	ALA M	4.544	46.429	-18.635
	74	ALA CA	7.478	47.591	-18.959	74	ALA C	7.740	47.648	-20.342
	74	ALA B	7.959	46.640	-21.054	74	ALA CB	8.653	47.446	-17.925
	75	LEU M	7.650	48.704	-21.839	75	LEU CA	7.812	48.968	-22.456
	75	LEU C	9.192	48.568	-22.966	75	LEU B	10.162	48.750	-22.253
	75	LEU CB	7.548	50.471	-22.809	75	LEU CG	4.123	50.913	-22.379
10	75	LEU CD1	6.079	52.436	-22.380	75	LEU CD2	5.094	50.442	-23.405
	76	ASM M	9.167	48.103	-24.169	76	ASM MD2	12.385	46.432	-26.304
	76	ASM OD1	10.950	45.840	-27.928	76	ASM CG	11.195	46.274	-26.802
	76	ASM CB	10.810	46.651	-25.988	76	ASM CA	10.359	47.738	-24.938
	76	ASM C	10.783	49.048	-25.643	76	ASM B	10.157	49.479	-26.619
	77	ASM B	11.004	49.664	-25.071	77	ASM CA	12.220	50.957	-25.681
	77	ASM C	13.797	51.029	-25.348	77	ASM B	14.364	49.979	-25.313
	77	ASM CB	11.335	52.076	-25.117	77	ASM CG	11.250	52.027	-23.616
15	77	ASM OD1	12.032	51.346	-22.917	77	ASM MD2	10.294	52.741	-23.025
	78	SER M	14.125	52.267	-25.164	78	SER CA	15.513	52.614	-24.986
	78	SER C	15.810	52.742	-23.436	78	SER B	16.982	53.071	-23.164
	78	SER CB	15.985	53.941	-25.587	78	SER CG	15.926	53.870	-24.999
	79	ILE M	14.858	52.565	-22.529	79	ILE CA	15.155	52.784	-21.120
	79	ILE C	14.617	51.683	-20.230	79	ILE B	13.843	50.841	-28.679
	79	ILE CB	14.471	54.174	-20.697	79	ILE CG1	12.945	54.832	-28.814
	79	ILE CG2	14.997	55.320	-21.612	79	ILE CD1	12.135	55.176	-28.155
20	80	GLY M	14.995	51.760	-18.981	80	GLY CA	14.476	50.948	-17.913
	80	GLY C	14.612	49.448	-18.219	80	GLY B	15.719	48.994	-18.544
	81	VAL M	13.513	48.766	-17.980	81	VAL CA	13.411	47.286	-18.061
	81	VAL C	12.511	46.919	-19.217	81	VAL B	12.260	47.739	-20.117
	81	VAL CB	13.001	46.755	-16.677	81	VAL CG1	14.030	47.084	-15.573
	81	VAL CG2	11.638	47.261	-16.231	82	LEU M	12.126	45.645	-19.216
	82	LEU CA	11.312	45.820	-20.256	82	LEU C	10.390	44.028	-19.510
	82	LEU B	10.858	43.356	-18.600	82	LEU CB	12.204	44.219	-21.229
25	82	LEU CG	11.430	43.568	-22.366	82	LEU CD1	10.796	44.657	-23.223
	82	LEU CD2	12.359	42.675	-23.192	83	GLY M	9.131	44.180	-19.816
	83	GLY CA	8.133	43.321	-19.114	83	GLY C	8.027	42.011	-19.925
	83	GLY B	8.546	41.822	-21.826	84	VAL M	7.272	41.112	-19.283
	84	VAL CA	6.973	39.807	-19.888	84	VAL C	6.164	40.830	-21.140
	84	VAL B	6.424	39.472	-22.196	84	VAL CB	6.256	38.920	-18.841
	84	VAL CG1	5.680	37.677	-19.557	84	VAL CG2	7.190	38.507	-17.705
30	85	ALA M	5.154	40.924	-21.024	85	ALA CA	4.217	41.194	-22.150
	85	ALA C	4.213	42.683	-22.396	85	ALA B	3.268	43.401	-22.030
	85	ALA CB	2.846	40.663	-21.748	86	PRO M	5.240	43.106	-23.059
	86	PRO CA	5.413	46.635	-23.285	86	PRO C	4.321	45.371	-23.947
	86	PRO B	4.291	46.605	-23.849	86	PRO CB	4.822	44.784	-23.813
	86	PRO CG	7.830	43.466	-24.546	86	PRO CD	4.377	42.440	-23.636
	87	SER M	3.548	44.676	-24.769	87	SER CA	2.489	45.324	-25.529
	87	SER C	1.303	45.132	-24.897	87	SER B	8.162	45.513	-25.619
35	87	SER CB	2.401	46.777	-26.927	87	SER CG	3.591	45.143	-27.583
	88	ALA M	1.017	44.564	-23.742	88	ALA CB	-0.163	43.518	-21.828
	88	ALA CA	-0.273	44.353	-23.084	88	ALA C	-0.090	45.717	-22.690
	88	ALA B	-0.174	46.717	-22.435	89	SER M	-2.219	45.691	-22.678
	89	SER CG	-4.146	47.102	-24.288	89	SER CA	-4.343	46.983	-22.898
	89	SER CB	-3.001	46.867	-22.227	89	SER C	-3.136	46.788	-28.727
	89	SER B	-3.793	45.864	-20.209	90	LEU M	-2.446	47.656	-28.037
40	90	LEU CA	-2.378	47.667	-18.593	90	LEU C	-3.483	48.438	-17.864
	90	LEU B	-3.582	49.604	-18.215	90	LEU CB	-0.951	48.273	-18.426
	90	LEU CG	-0.233	47.851	-17.174	90	LEU CD1	-0.026	46.361	-17.219
	90	LEU CD2	1.160	48.524	-17.847	91	TYR M	-4.264	47.964	-16.938
	91	TYR CA	-5.258	48.678	-16.137	91	TYR C	-4.873	48.750	-14.685

45

50

55

5	91	TYR D	-4.494	47.749	-14.073	91	TYR CB	-6.686	48.093	-14.314
	91	TYR CL	-7.094	48.237	-17.741	91	TYR CB1	-6.595	47.415	-18.755
	91	TYR CD2	-7.971	49.275	-18.149	91	TYR CE1	-6.905	47.572	-20.090
	91	TYR CE2	-8.315	49.421	-19.492	91	TYR CZ	-7.794	48.502	-20.463
	91	TYR DM	-8.202	48.752	-21.764	92	ALA M	-6.895	49.958	-14.104
10	92	ALA CA	-6.549	50.199	-12.707	92	ALA C	-5.823	50.033	-11.903
	92	ALA D	-6.723	50.090	-12.850	92	ALA CB	-3.997	51.621	-12.480
	93	VAL M	-5.959	48.993	-11.129	93	VAL CA	-7.103	48.054	-10.325
	93	VAL C	-6.700	49.014	-8.899	93	VAL D	-6.181	47.993	-8.372
	93	VAL CB	-7.957	47.555	-10.631	93	VAL CG1	-0.213	47.488	-9.725
15	93	VAL CG2	-0.195	47.370	-12.072	94	LVS M	-6.907	50.217	-8.327
	94	LVS CA	-6.370	50.464	-6.999	94	LVS C	-7.331	49.905	-5.094
	94	LVS D	-8.458	50.480	-5.783	94	LVS CB	-6.051	51.976	-6.818
	94	LVS CG	-5.394	52.320	-5.467	94	LVS CD	-6.868	53.785	-5.582
	94	LVS CE	-4.399	54.200	-4.199	94	LVS M2	-3.735	55.544	-4.387
20	95	VAL M	-6.909	49.071	-5.026	95	VAL CA	-7.646	48.457	-3.920
	95	VAL C	-6.919	48.499	-2.568	95	VAL D	-7.425	48.156	-1.501
	95	VAL CB	-8.104	47.030	-4.319	95	VAL CG1	-0.848	46.852	-5.619
	95	VAL CG2	-6.900	46.100	-4.332	96	LEU M	-5.676	48.974	-2.604
	96	LEU CA	-4.782	49.103	-1.486	96	LEU C	-6.331	50.559	-1.321
25	96	LEU D	-3.942	51.121	-2.336	96	LEU CB	-3.509	48.241	-1.573
	96	LEU CG	-3.593	46.799	-2.072	96	LEU CD1	-2.207	46.184	-2.163
	96	LEU CD2	-4.489	46.082	-1.045	97	GLY M	-4.326	50.975	-0.086
	97	GLY CA	-3.890	52.307	0.287	97	GLY C	-2.363	52.437	0.385
	97	GLY D	-1.619	51.463	0.165	98	ALA M	-1.954	53.648	0.750
30	98	ALA CB	-0.420	55.478	1.510	98	ALA CA	-0.563	54.068	0.965
	98	ALA C	0.108	53.110	1.917	98	ALA D	1.393	52.921	1.663
	99	ASP M	-0.504	52.573	2.912	99	ASP CD2	-2.631	51.042	6.151
	99	ASP CD1	-2.730	50.902	4.003	99	ASP CG	-2.003	51.131	5.048
	99	ASP CB	-0.640	51.603	5.175	99	ASP CA	0.101	51.610	3.055
35	99	ASP C	0.146	50.165	3.320	99	ASP D	0.735	49.313	4.029
	100	GLY M	-0.424	49.803	2.168	100	GLY CA	-0.343	48.521	1.615
	100	GLY C	-1.520	47.651	2.002	100	GLY D	-1.649	46.512	1.479
	101	SER M	-2.342	48.120	2.900	101	SER CA	-3.542	47.380	3.315
	101	SER C	-4.759	47.094	2.532	101	SER D	-4.758	48.972	1.907
40	101	SER CB	-3.716	47.447	4.817	101	SER DC	-4.411	48.434	5.209
	102	GLY M	-5.021	47.092	2.577	102	GLY CA	-7.877	47.422	1.894
	102	GLY C	-8.166	46.534	2.528	102	GLY D	-7.800	45.431	3.030
	103	GLN M	-9.377	47.058	2.498	103	GLN CA	-10.535	46.297	3.020
	103	GLN C	-10.963	45.232	2.022	103	GLN	-10.779	45.402	0.017
45	103	GLN CB	-11.671	47.307	3.274	103	GLN CG	-11.348	48.005	4.586
	103	GLN CD	-12.360	49.104	4.915	103	GLN DE1	-12.159	49.016	5.902
	103	GLN ME2	-13.419	49.197	6.112	104	TYR M	-11.611	46.141	2.451
	104	TYR CA	-12.068	43.124	1.506	104	TYR C	-13.031	43.690	0.473
	104	TYR D	-12.939	43.276	-0.607	104	TYR CB	-12.697	41.866	2.143
50	104	TYR CG	-11.629	40.829	2.472	104	TYR CD1	-11.019	39.709	3.377
	104	TYR CD2	-10.379	40.959	1.060	104	TYR CE1	-10.805	38.885	3.707
	104	TYR CE2	-9.352	40.057	2.171	104	TYR CZ	-9.564	39.022	3.081
	104	TYR DM	-8.401	38.191	3.324	105	SER M	-13.909	44.572	8.903
	105	SER CA	-14.077	45.166	-0.034	105	SER C	-14.172	45.920	-1.159
55	105	SER D	-14.759	45.935	-2.258	105	SER CB	-15.080	46.121	0.601
	105	SER DC	-15.209	47.039	1.450	106	TRP M	-13.079	46.625	-0.834
	106	TRP CA	-12.421	47.391	-1.948	106	TRP C	-11.895	46.436	-3.012
	106	TRP D	-12.021	46.640	-4.245	106	TRP CB	-11.321	48.254	-1.355
	106	TRP CG	-11.645	49.111	-0.206	106	TRP CD1	-12.062	49.524	0.264
60	106	TRP CD2	-10.450	49.812	0.501	106	TRP ME1	-12.491	50.350	1.360
	106	TRP CE2	-11.359	50.573	1.561	106	TRP CE3	-9.275	49.852	0.576
	106	TRP CZ	-10.671	51.310	2.500	106	TRP CZ3	-0.560	50.563	1.525
	106	TRP CM2	-9.293	51.291	2.455	107	ILE M	-11.339	45.330	-2.481
	107	ILE CA	-10.765	44.250	-3.325	107	ILE C	-11.955	43.594	-4.190
65	107	ILE D	-11.695	43.474	-5.398	107	ILE CB	-9.944	43.103	-2.523
	107	ILE CG1	-8.634	43.784	-1.936	107	ILE CG2	-9.632	41.930	-3.381
	107	ILE CD1	-8.283	42.998	-0.627	109	ILE M	-12.094	43.292	-3.577

100	ILE CA	-14.114	42.722	-4.321	108	ILE C	-14.639	43.694	-5.384
100	ILE D	-14.894	43.329	-4.552	108	ILE C0	-15.246	42.263	-3.320
100	ILE CG1	-14.726	41.077	-2.482	108	ILE CG2	-16.940	42.824	-4.093
100	ILE CD1	-15.452	40.845	-1.131	109	ASM M	-14.751	44.958	-4.981
109	ASM CA	-15.204	46.018	-5.916	109	ASM C	-14.232	46.067	-7.084
109	ASM D	-14.668	46.272	-8.235	109	ASM C0	-15.200	47.359	-5.207
109	ASM CG	-16.328	47.686	-4.353	109	ASM CD1	-17.653	46.695	-4.646
109	ASM MD2	-14.633	48.447	-3.442	110	GLV M	-12.951	45.908	-4.774
110	GLV CA	-11.952	45.917	-7.863	110	GLV C	-12.108	44.712	-8.812
110	GLV D	-11.929	44.929	-10.834	111	ILE M	-12.379	43.539	-8.246
111	ILE CA	-12.603	42.334	-9.099	111	ILE C	-13.859	42.560	-9.942
111	ILE D	-13.721	42.304	-11.148	111	ILE C0	-12.734	40.948	-8.364
111	ILE CG1	-12.421	40.501	-7.655	111	ILE CG2	-13.122	39.791	-9.347
111	ILE CD1	-11.588	39.706	-6.336	112	GLU M	-14.893	43.075	-9.280
112	GLU CA	-16.118	43.376	-10.046	112	GLU C	-15.872	44.347	-11.171
112	GLU D	-16.467	44.130	-12.746	112	GLU C0	-17.229	43.899	-9.141
112	GLU CG	-17.847	42.917	-8.135	112	GLU CD	-18.724	41.824	-8.685
112	GLU DE1	-19.041	40.864	-9.016	112	GLU DE2	-19.123	41.928	-9.866
113	TRP M	-15.094	45.403	-10.971	113	TRP CA	-14.756	46.400	-12.000
113	TRP C	-14.876	45.663	-13.140	113	TRP D	-14.319	45.932	-14.332
113	TRP C0	-13.882	47.553	-11.434	113	TRP CG	-13.486	48.556	-12.481
113	TRP CD1	-14.148	49.736	-12.681	113	TRP CD2	-12.441	48.552	-13.463
113	TRP ME1	-13.597	50.443	-13.723	113	TRP CE2	-12.545	49.761	-14.215
113	TRP CE3	-11.651	47.645	-13.809	113	TRP CZ2	-11.696	50.045	-15.274
113	TRP CZ3	-10.610	47.899	-14.879	113	TRP CM2	-10.752	49.074	-15.603
114	ALA M	-13.089	44.801	-12.832	114	ALA CA	-12.333	44.065	-13.874
114	ALA C	-13.199	43.179	-14.752	114	ALA D	-12.963	43.074	-15.978
114	ALA C0	-11.299	43.192	-13.140	115	ILE M	-14.174	42.540	-14.119
115	ILE CA	-15.870	41.640	-14.897	115	ILE C	-15.928	42.485	-15.856
115	ILE D	-16.077	42.225	-17.070	115	ILE C0	-16.080	40.840	-13.922
115	ILE CG1	-15.218	39.836	-13.843	115	ILE CG2	-17.151	40.168	-14.755
115	ILE CD1	-16.004	39.411	-11.743	116	ALA M	-16.534	43.527	-15.267
116	ALA CA	-17.390	44.440	-16.050	116	ALA C	-16.706	45.069	-17.278
116	ALA D	-17.323	45.255	-18.343	116	ALA C0	-18.011	45.510	-15.151
117	ASM M	-15.423	45.390	-17.122	117	ASM CA	-14.553	45.967	-18.139
117	ASM C	-13.827	44.974	-19.034	117	ASM D	-12.997	45.436	-19.820
117	ASM C0	-13.615	46.958	-17.426	117	ASM CG	-14.480	48.177	-16.939
117	ASM CD1	-14.565	49.082	-17.773	117	ASM MD2	-14.931	48.249	-15.734
118	ASM M	-14.223	43.725	-10.967	118	ASM CA	-13.740	42.642	-19.832
118	ASM C	-12.240	42.444	-19.843	118	ASM D	-11.617	42.309	-20.932
118	ASM C0	-14.247	42.863	-21.279	118	ASM CG	-15.737	43.060	-21.395
118	ASM CD1	-14.510	42.321	-20.759	118	ASM MD2	-16.136	44.096	-22.133
119	MET M	-11.686	42.500	-18.675	119	MET CA	-10.232	42.222	-18.478
119	MET C	-10.025	40.734	-18.928	119	MET D	-10.888	39.838	-18.759
119	MET C0	-9.810	42.461	-17.055	119	MET CG	-9.880	43.883	-16.582
119	MET SD	-8.788	44.943	-17.526	119	MET CE	-9.982	44.061	-18.263
120	ASP M	-8.904	40.437	-19.584	120	ASP CA	-8.480	39.110	-20.830
120	ASP C	-7.822	34.390	-18.856	120	ASP D	-8.838	37.189	-18.690
120	ASP C0	-7.555	39.154	-21.236	120	ASP CG	-8.237	39.730	-22.454
120	ASP CD1	-7.801	40.706	-23.084	120	ASP DD2	-9.327	39.135	-22.739
121	VAL M	-7.071	39.117	-18.115	121	VAL CA	-6.224	38.601	-16.974
121	VAL C	-6.296	39.534	-15.786	121	VAL D	-6.284	40.708	-15.909
121	VAL C0	-4.715	38.507	-17.496	121	VAL CG1	-3.758	38.176	-16.427
121	VAL CG2	-4.787	37.916	-18.846	122	ILE M	-6.318	38.978	-14.598
122	ILE CA	-6.248	39.799	-13.397	122	ILE C	-5.820	39.262	-12.627
122	ILE D	-4.829	38.812	-12.469	122	ILE C0	-7.476	39.604	-12.666
122	ILE CG1	-8.686	40.392	-13.063	122	ILE CG2	-7.221	39.883	-10.954
122	ILE CD1	-9.976	39.788	-12.383	123	ASM M	-4.263	40.222	-12.118
123	ASM CA	-3.145	39.854	-11.232	123	ASM C	-3.502	40.404	-9.861
123	ASM D	-3.708	41.631	-9.833	123	ASM C0	-1.828	40.478	-11.697
123	ASM CG	-8.692	40.048	-10.777	123	ASM CD1	-8.063	38.990	-11.018
123	ASM MD2	-8.366	40.747	-9.728	124	MET M	-3.458	39.604	-8.832
124	MET CA	-3.650	39.973	-7.438	124	MET C	-2.423	39.603	-6.614

45

50

55

5	124	MIT D	-2.304	38.508	-6.893	124	MIT C8	-4.943	39.387	-6.893
	124	MIT CG	-6.138	40.982	-7.673	124	MIT SC	-7.983	39.472	-6.138
	124	MIT C1	-7.969	38.093	-7.962	125	SEP M	-1.484	40.496	-6.987
	125	SEP CA	-0.193	40.287	-3.769	125	SEP C	-0.422	40.712	-4.324
	125	SEP D	0.239	41.617	-3.803	125	SEP C8	1.021	41.027	-4.324
	125	SEP DG	1.444	40.496	-7.575	126	LEU M	-1.433	40.878	-3.773
	126	LEU CA	-1.842	40.347	-2.386	126	LEU C	-2.438	39.896	-1.807
	126	LEU D	-2.844	38.136	-2.529	126	LEU C8	-2.791	41.868	-2.410
	126	LEU CG	-3.988	41.447	-3.333	126	LEU CG1	-3.278	41.131	-2.578
	126	LEU CG2	-4.179	42.760	-4.873	127	GLY M	-2.522	39.882	-0.681
	127	GLY CA	-3.835	37.871	0.193	127	GLY C	-3.176	38.180	1.682
	127	GLY D	-2.446	39.830	2.220	128	GLY M	-4.121	37.443	2.222
	128	GLY CA	-4.475	37.496	3.642	128	GLY C	-6.644	36.838	4.184
	128	GLY D	-4.983	38.158	3.276	129	PRO M	-4.519	38.857	8.482
10	129	PRO CA	-4.671	34.523	8.998	129	PRO C	-4.116	34.886	4.082
	129	PRO D	-6.338	32.887	6.303	129	PRO C8	-4.860	34.484	7.384
	129	PRO CG	-4.419	36.116	7.727	129	PRO CG	-4.239	36.870	6.618
	130	SEP M	-7.051	33.813	8.912	130	SEP CA	-8.470	34.611	4.023
	130	SEP C	-0.218	34.884	4.726	130	SEP D	-8.949	33.881	4.829
	130	SEP C8	-9.849	31.331	7.216	130	SEP DG	-8.723	34.624	8.483
	131	GLY M	-10.883	33.967	4.349	131	GLY CA	-10.824	34.229	3.874
	131	GLY C	-12.205	34.713	3.342	131	GLY D	-12.495	34.722	4.751
15	132	SEP M	-13.840	31.958	2.594	132	SEP CA	-14.607	35.433	3.811
	132	SEP C	-13.289	34.805	1.936	132	SEP D	-14.798	34.886	8.824
	132	SEP C8	-14.580	36.927	3.145	132	SEP DG	-14.893	37.939	1.875
	133	ALA M	-16.547	34.588	2.294	133	ALA CA	-17.507	34.857	1.324
	133	ALA C	-17.630	34.965	8.897	133	ALA D	-17.743	34.427	-1.816
	133	ALA C8	-18.866	33.828	1.996	134	ALA M	-17.683	36.288	8.294
	134	ALA CA	-17.872	37.219	-0.792	134	ALA C	-18.635	37.369	-1.674
	134	ALA D	-16.781	37.585	-2.849	134	ALA C8	-18.263	38.600	-0.187
20	135	LEU M	-18.478	37.229	-1.846	135	LEU CA	-14.197	37.244	-1.884
	135	LEU C	-16.138	36.903	-2.705	135	LEU D	-13.794	34.820	-1.890
	135	LEU C8	-13.038	37.328	-0.798	135	LEU CG	-11.693	37.180	-1.588
	135	LEU CG1	-11.460	38.415	-2.212	135	LEU CG2	-10.582	36.807	-0.519
	136	LYS M	-14.809	34.823	-2.173	136	LYS CA	-14.543	33.597	-3.813
	136	LYS C	-13.544	33.739	-4.150	136	LYS C	-18.279	33.431	-3.385
	136	LYS C8	-14.903	32.341	-2.186	136	LYS CG	-14.743	31.867	-3.943
	136	LYS CG	-18.883	29.892	-2.134	136	LYS C8	-15.743	28.707	-2.778
25	136	LYS M2	-18.308	28.411	-4.160	137	ALA M	-14.764	34.260	-3.847
	137	ALA CA	-17.793	34.416	-4.883	137	ALA C	-17.338	36.303	-6.045
	137	ALA D	-17.705	35.049	-7.208	137	ALA C8	-18.094	34.941	-4.243
	138	ALA M	-16.329	36.301	-3.729	138	ALA C4	-16.891	37.311	-6.685
	138	ALA C	-16.953	36.696	-7.857	138	ALA D	-14.985	36.843	-8.742
	138	ALA C8	-15.122	38.867	-5.934	139	VAL M	-13.950	33.959	-7.827
	139	VAL CA	-12.946	33.291	-7.837	139	VAL C	-13.623	34.220	-8.720
	139	VAL D	-13.108	34.870	-9.877	139	VAL C8	-11.830	34.671	-6.968
30	139	VAL CG1	-10.919	33.856	-7.866	139	VAL CG2	-11.898	33.780	-6.213
	140	ASP M	-16.553	33.536	-8.122	140	ASP CA	-15.274	32.494	-8.929
	140	ASP C	-16.923	33.131	-10.084	140	ASP D	-16.980	32.579	-11.190
	140	ASP C8	-16.149	31.549	-8.138	140	ASP CG	-15.388	38.640	-7.186
	140	ASP CG1	-16.178	30.483	-7.282	140	ASP CG2	-16.139	38.132	-4.329
	141	LYS M	-16.658	34.263	-9.810	141	LYS CA	-17.373	31.806	-18.568
	141	LYS C	-18.373	35.418	-11.946	141	LYS D	-16.780	33.248	-13.111
	141	LYS C8	-18.839	36.275	-10.321	141	LYS CG	-18.884	37.834	-11.386
35	141	LYS CG	-18.586	38.187	-10.536	141	LYS CG	-20.572	39.851	-11.250
	141	LYS M2	-21.130	40.837	-10.273	142	ALA M	-15.167	33.848	-11.966
	142	ALA CA	-14.173	36.192	-12.616	142	ALA C	-13.818	33.810	-13.871
	142	ALA D	-13.770	33.169	-14.755	142	ALA C8	-12.870	36.697	-11.948
	143	VAL M	-13.582	33.886	-12.832	143	VAL CA	-13.168	32.785	-13.650
	143	VAL C	-14.346	32.233	-14.496	143	VAL D	-16.149	31.886	-15.639
	143	VAL C8	-12.551	31.473	-12.716	143	VAL CG1	-12.280	38.370	-13.461
	143	VAL CG2	-13.395	32.193	-12.814	144	ALA M	-15.531	32.238	-13.875
40	144	ALA CA	-16.744	31.834	-14.641	144	ALA C	-16.928	32.681	-15.861

45

50

55

5	146	ALA C	-17.380	32.243	-16.950	146	ALA CB	-17.942	31.968	-13.788
	148	SIF M	-16.307	33.948	-13.701	148	SIF CA	-16.682	34.017	-16.786
	149	SIF C	-15.609	34.773	-17.829	149	SIF D	-15.918	35.321	-18.893
	149	SIF CB	-17.816	34.376	-16.614	149	SIF DG	-15.882	36.988	-19.849
	146	GLV M	-14.977	33.986	-17.565	146	GLV CA	-13.619	33.799	-18.671
10	146	GLV C	-12.173	34.491	-12.385	146	GLV D	-11.420	34.386	-19.266
	147	VAL M	-12.150	35.162	-17.254	147	VAL CA	-19.874	39.836	-16.912
	147	VAL C	-9.850	34.834	-16.323	147	VAL D	-18.171	39.991	-15.486
	147	VAL CB	-11.152	34.977	-15.839	147	VAL CG1	-9.894	37.803	-15.878
	147	VAL CG2	-12.340	37.913	-16.230	148	VAL M	-8.583	35.818	-16.683
15	148	VAL CA	-7.682	34.230	-16.808	148	VAL C	-7.137	34.907	-14.791
	148	VAL D	-4.840	36.133	-14.790	148	VAL CB	-6.273	34.126	-16.938
	148	VAL CG1	-5.079	33.483	-16.281	148	VAL CG2	-6.398	33.632	-18.262
	149	VAL M	-7.258	34.355	-13.531	149	VAL CA	-6.987	34.963	-12.249
	149	VAL C	-8.769	34.385	-11.613	149	VAL D	-5.624	33.173	-11.639
20	149	VAL CB	-8.224	34.890	-11.313	149	VAL CG1	-7.893	35.619	-19.089
	149	VAL CG2	-9.456	35.386	-12.094	150	VAL M	-6.732	35.301	-11.484
	150	VAL CA	-3.393	34.917	-19.901	150	VAL C	-3.157	35.623	-9.839
	150	VAL D	-3.892	36.778	-9.480	150	VAL CB	-2.274	35.393	-11.951
	150	VAL CG1	-8.973	34.633	-11.461	150	VAL CG2	-2.678	34.843	-13.381
25	151	ALA M	-2.568	34.946	-8.393	151	ALA CA	-2.341	35.382	-7.287
	151	ALA C	-1.980	35.034	-6.637	151	ALA D	-0.618	33.889	-6.984
	151	ALA CB	-3.597	35.390	-6.307	152	ALA M	-0.490	35.987	-5.822
	152	ALA CA	0.714	35.438	-5.112	152	ALA C	0.384	34.320	-4.188
	152	ALA D	-8.718	34.466	-3.467	152	ALA CB	1.266	36.687	-4.294
30	153	ALA M	1.125	33.302	-3.912	153	ALA CA	0.840	32.258	-2.943
	153	ALA C	0.931	32.725	-3.811	153	ALA D	0.317	32.192	-0.999
	153	ALA CB	1.750	31.038	-3.193	154	GLV M	1.827	33.693	-1.244
	154	GLV CA	2.043	34.211	0.125	154	GLV C	3.319	34.849	0.550
	154	GLV D	4.189	33.267	-8.118	155	ASM M	3.958	34.788	1.568
35	155	ASM CA	5.344	34.787	2.837	155	ASM C	5.399	34.258	3.462
	155	ASM D	6.101	34.829	4.293	155	ASM CB	6.008	36.158	1.984
	155	ASM CG	5.890	36.702	0.300	155	ASM CG1	6.123	36.863	-8.834
	155	ASM CG2	5.454	37.963	0.352	156	GLU M	4.711	33.168	3.673
	156	GLU CA	4.633	32.937	4.970	156	GLU C	5.322	31.328	0.163
40	156	GLU D	5.374	30.637	6.222	156	GLU CB	3.203	31.980	0.189
	156	GLU CG	2.491	32.442	0.268	156	GLU CD	2.394	33.951	6.270
	156	GLU DE1	1.744	34.322	9.312	156	GLU DE2	3.186	34.656	7.146
	157	GLV M	6.339	31.057	4.227	157	GLV CA	7.306	29.917	4.387
	157	GLV C	6.503	28.622	4.533	157	GLV D	5.416	28.348	4.889
45	158	TMR M	7.147	27.793	5.382	158	TMR CG2	8.079	29.396	3.839
	158	TMR CG1	8.787	25.487	6.217	158	TMR CB	7.564	25.346	3.296
	158	TMR CA	6.352	26.487	5.752	159	TMR C	6.188	26.420	7.137
	158	TMR D	6.479	27.335	7.977	159	SER M	5.338	25.441	7.497
	159	SER DG	3.141	25.904	18.525	159	SER CB	3.673	26.189	9.212
50	159	SER CA	4.838	25.210	8.955	159	SER C	4.494	23.728	8.944
	159	SER D	3.339	23.281	9.030	160	GLV M	5.974	22.967	8.839
	160	GLV CA	5.434	21.504	8.893	160	GLV C	4.576	21.849	7.738
	160	GLV D	4.808	21.324	6.955	161	SER M	3.923	20.318	8.136
	161	SER CA	2.654	19.777	7.854	161	SER C	1.477	20.788	6.766
55	161	SER D	9.696	20.347	5.869	161	SER CB	2.344	18.293	7.271
	161	SER DG	1.834	18.029	8.585	162	SER M	1.393	21.841	7.659
	162	SER CA	0.167	22.721	7.113	162	SER C	0.430	23.553	5.848
	162	SER D	1.333	23.040	5.394	162	SER CB	-0.219	23.666	8.242
	162	SER DG	8.184	23.891	9.480	163	SER M	-0.678	23.923	5.197
60	163	SER CA	-0.611	24.750	3.992	163	SER C	-0.441	24.177	4.533
	163	SER D	-1.878	24.548	5.304	163	SER CB	-1.890	24.642	3.211
	163	SER DG	-1.992	23.718	7.331	164	TMR M	0.387	24.932	3.932
	164	TMR CA	0.689	29.340	4.312	164	TMR C	0.185	29.286	3.194
	164	TMR D	0.483	30.502	3.278	164	TMR CB	2.099	28.518	4.818
65	164	TMR CG1	2.984	28.282	3.692	164	TMR CG2	2.397	27.618	6.081
	165	VAL M	-8.513	28.742	2.190	165	VAL CA	-8.959	29.342	1.818
70	165	VAL C	-2.928	28.548	1.697	165	VAL D	-2.929	30.192	2.288

5	165	VAL	CD	-1.339	28.024	-0.161	165	VAL	CD1	-1.047	29.357	-1.374
	165	VAL	CD2	-3.210	27.716	-0.099	166	GLY	M	-1.910	31.071	1.129
	166	GLY	CA	-2.943	32.778	3.020	166	GLY	C	-4.090	32.810	0.617
	166	GLY	D	-4.124	32.184	-0.396	167	TYR	M	-9.014	33.730	0.970
	167	TYR	CA	-6.223	34.066	0.113	167	TYR	C	-9.993	31.309	-0.606
	167	TYR	D	-8.474	36.203	0.064	167	TYR	CD	-7.464	34.232	0.964
	167	TYR	CG	-7.791	32.054	1.709	167	TYR	CD1	-7.200	32.703	2.047
	167	TYR	CD2	-8.710	32.116	1.133	167	TYR	CD1	-7.567	31.520	3.613
	167	TYR	CD2	-9.068	30.933	1.009	167	TYR	CD2	-8.486	30.671	3.046
	167	TYR	DM	-8.080	29.481	3.658	168	PRO	M	-6.310	31.409	-1.030
	168	PRO	CG	-6.943	36.376	-3.938	168	PRO	CD	-6.273	34.752	-2.624
	168	PRO	CG	-7.964	31.344	-3.303	168	PRO	CA	-7.134	34.457	-2.560
10	168	PRO	C	-6.390	33.336	-3.270	168	PRO	D	-7.097	32.520	-3.912
	169	GLY	M	-3.086	33.193	-3.109	169	GLY	CA	-4.446	32.077	-3.927
	169	GLY	C	-4.937	30.702	-3.470	169	GLY	D	-4.000	29.733	-4.249
	170	LVS	M	-3.402	30.979	-2.255	170	LVS	CA	-8.016	29.268	-1.745
	170	LVS	C	-7.033	28.773	-2.316	170	LVS	D	-7.398	27.554	-2.524
	170	LVS	CG	-4.246	29.294	-0.306	170	LVS	CG	-5.795	28.106	0.983
	170	LVS	CD	-4.230	28.289	2.031	170	LVS	CD	-5.731	27.271	3.029
	170	LVS	CD2	-4.239	27.463	3.213	171	TYR	M	-7.030	29.616	-3.148
15	171	TYR	CA	-9.012	29.043	-3.059	171	TYR	C	-8.603	28.309	-0.113
	171	TYR	D	-7.760	28.714	-5.928	171	TYR	CD	-9.962	30.224	-4.242
	171	TYR	CG	-10.497	30.964	-3.047	171	TYR	CD1	-11.040	30.303	-1.982
	171	TYR	CD2	-10.456	32.374	-3.026	171	TYR	CD2	-11.920	31.003	-0.867
	171	TYR	CD2	-10.941	33.003	-1.936	171	TYR	CD2	-11.920	32.393	-0.006
	171	TYR	DM	-12.006	33.119	0.170	172	PRO	M	-9.297	27.204	-3.374
	172	PRO	CA	-9.093	26.417	-6.396	172	PRO	C	-9.233	27.156	-7.909
	172	PRO	D	-8.325	26.784	-0.801	172	PRO	CG	-10.167	28.329	-6.513
20	172	PRO	CG	-10.670	28.271	-0.096	172	PRO	CD	-10.364	26.469	-4.816
	173	SER	M	-10.017	28.167	-0.019	173	SER	CA	-10.220	28.018	-9.330
	173	SER	C	-9.025	29.773	-9.593	173	SER	D	-8.966	30.233	-10.742
	173	SER	CG	-11.528	29.623	-9.491	173	SER	CD	-11.593	30.546	-0.406
	174	VAL	M	-8.162	29.944	-8.614	174	VAL	CA	-7.013	30.091	-1.033
	174	VAL	C	-9.754	30.131	-0.068	174	VAL	D	-5.012	29.152	-1.344
	174	VAL	CG	-6.899	31.775	-7.594	174	VAL	CG1	-5.794	32.837	-7.617
	174	VAL	CG2	-8.220	32.303	-7.323	175	ILE	M	-4.913	30.729	-0.031
25	175	ILE	CA	-3.369	30.184	-10.024	175	ILE	C	-2.714	30.734	-0.894
	175	ILE	D	-2.430	31.930	-8.933	175	ILE	CG	-2.933	30.524	-11.410
	175	ILE	CG1	-3.837	29.978	-12.924	175	ILE	CG2	-1.451	30.009	-11.312
	175	ILE	CG1	-3.692	30.529	-13.946	176	ALA	M	-2.220	30.020	-7.928
	176	ALA	CA	-1.335	30.517	-6.870	176	ALA	C	0.120	30.301	-7.310
	176	ALA	D	0.433	29.218	-7.030	176	ALA	CG	-1.039	29.830	-5.561
	177	VAL	M	0.044	31.410	-7.100	177	VAL	CA	2.261	31.834	-7.636
	177	VAL	C	3.223	31.693	-6.473	177	VAL	D	3.170	32.657	-5.721
30	177	VAL	CG	2.439	32.607	-8.763	177	VAL	CG1	3.042	32.667	-9.392
	177	VAL	CG2	1.374	32.352	-9.043	178	GLY	M	6.077	30.634	-6.390
	178	GLY	CA	3.160	30.703	-9.333	178	GLY	C	6.446	31.233	-6.074
	178	GLY	D	6.490	31.435	-7.266	179	ALA	M	7.012	31.447	-5.287
	179	ALA	CA	0.715	32.037	-9.859	179	ALA	C	9.939	31.099	-1.779
	179	ALA	C	10.198	30.481	-4.710	179	ALA	CG	9.023	33.231	-4.973
	180	VAL	M	10.639	31.162	-6.005	180	VAL	CA	11.970	30.482	-6.901
35	180	VAL	C	13.048	31.595	-7.171	180	VAL	D	12.712	32.691	-7.427
	180	VAL	CG	12.075	29.314	-0.166	180	VAL	CG1	12.271	28.291	-7.053
	180	VAL	CG2	11.675	30.120	-9.500	181	ASP	M	14.267	31.293	-6.000
	181	ASP	CA	13.431	32.100	-7.039	181	ASP	C	10.942	31.004	-5.462
	181	ASP	D	10.339	31.090	-9.292	181	ASP	CG	16.468	31.921	-5.014
	181	ASP	CG	17.120	30.934	-9.971	181	ASP	CD1	17.109	29.703	-6.972
	181	ASP	CD2	17.680	30.216	-4.087	182	SER	M	17.007	32.306	-5.047
40	182	SER	CA	17.622	32.214	-10.101	182	SER	C	20.193	30.017	-12.404
	182	SER	D	10.365	30.492	-11.670	182	SER	CG	10.678	33.313	-10.464
	182	SER	CG	10.016	34.161	-10.475	183	SER	M	10.258	30.042	-9.623
	183	SER	CA	10.716	28.045	-9.444	183	SER	C	17.981	27.614	-9.947
	183	SER	D	17.039	26.415	-9.397	183	SER	CG	10.256	20.323	-5.007

45

50

55

5	181	SEB	CG	25.589	25.615	-0.231	184	ASB	M	26.373	26.094	-0.632
	184	ASB	CA	25.144	27.317	-1.390	184	ASB	C	26.931	26.720	-0.197
	184	ASB	B	24.198	25.759	-0.997	184	ASB	CB	25.014	26.341	-10.722
	184	ASB	CG	24.993	26.998	-12.076	184	ASB	CD1	24.798	26.184	-12.277
	184	ASB	CD2	25.352	26.210	-13.976	185	GLM	M	25.942	27.247	-7.199
10	185	GLM	CA	25.274	26.646	-1.833	185	GLM	C	24.290	27.694	-5.203
	185	GLM	B	24.199	25.726	-1.334	185	GLM	CB	24.999	26.569	-5.191
	185	GLM	CG	26.539	26.242	-3.614	185	GLM	CD	25.011	26.182	-3.296
	185	GLM	CD1	25.864	25.799	-4.961	185	GLM	CD2	25.266	26.386	-1.936
	186	ABC	M	23.278	26.918	-4.648	186	ABC	CA	22.183	27.774	-3.841
15	186	ABC	C	22.780	25.782	-2.866	186	ABC	B	22.698	26.384	-2.893
	186	ABC	CB	21.213	26.843	-3.114	186	ABC	CG	20.214	27.471	-2.161
	186	ABC	CD	9.467	26.337	-1.668	186	ABC	ME	9.866	26.333	-0.117
	186	ABC	CI	9.941	26.879	1.839	186	ABC	MM1	9.367	27.888	1.658
	186	ABC	MM2	10.966	26.321	1.783	187	ALA	M	22.294	26.009	-2.893
20	187	ALA	CA	22.723	31.864	-1.895	187	ALA	C	22.262	30.884	-0.917
	187	ALA	B	21.151	30.843	-0.387	187	ALA	CB	22.144	32.482	-2.344
	188	SEB	M	23.031	30.770	0.549	188	SEB	CA	22.671	30.286	1.868
	188	SEB	C	21.316	30.847	2.412	188	SEB	B	20.740	30.111	3.312
	188	SEB	CB	23.767	30.456	2.928	188	SEB	CG	24.137	31.826	2.841
25	189	PME	M	20.943	32.010	1.974	189	PME	CA	9.697	32.688	2.418
	189	PME	C	8.699	32.198	1.609	189	PME	B	7.389	32.956	2.011
	189	PME	CB	9.787	34.217	2.243	189	PME	CG	10.117	34.698	0.867
	189	PME	CD1	9.147	34.830	-0.121	189	PME	CD2	11.419	35.116	0.867
	189	PME	CD1	9.483	33.187	-1.411	189	PME	CD2	11.769	35.343	-0.701
30	189	PME	CI	18.786	35.856	-1.728	190	SEB	M	8.703	31.526	0.499
	190	SEB	CA	7.626	31.996	-0.391	190	SEB	C	6.663	30.162	0.328
	190	SEB	B	7.834	29.083	0.866	190	SEB	CB	0.181	30.590	-1.788
	190	SEB	CG	7.136	30.337	-2.618	191	SEB	M	9.308	30.951	0.326
	191	SEB	CA	4.341	29.696	0.987	191	SEB	C	4.261	28.330	0.223
35	191	SEB	B	4.943	28.269	-0.995	191	SEB	CB	3.815	30.411	0.911
	191	SEB	CG	2.729	31.291	1.954	192	VAL	M	3.786	27.318	0.928
	192	VAL	CA	3.629	25.932	0.391	192	VAL	C	2.284	25.291	0.686
	192	VAL	B	1.959	25.698	1.198	192	VAL	CB	4.781	25.127	1.988
	192	VAL	CG1	6.144	25.727	0.722	192	VAL	CG2	4.617	25.104	1.592
40	193	GLY	M	1.938	24.172	0.847	193	GLY	CA	8.629	25.564	0.418
	193	GLY	C	0.081	23.029	-0.901	193	GLY	B	8.530	25.244	-2.815
	194	PRC	M	-1.023	22.281	-0.722	194	PRC	CA	-1.662	21.851	-1.873
	194	PRC	C	-2.237	22.605	-2.914	194	PRC	B	-2.403	22.244	-6.883
	194	PRC	CB	-2.769	20.783	-1.210	194	PRC	CG	-2.311	20.623	0.213
45	194	PRC	CD	-1.633	21.954	0.578	195	GLU	M	-2.922	23.793	-2.439
	195	GLU	CA	-3.145	24.930	-3.232	195	GLU	C	-2.093	23.631	-4.858
	195	GLU	B	-2.516	26.298	-0.936	195	GLU	CB	-4.843	25.784	-2.678
	195	GLU	CG	-4.942	25.134	-1.435	195	GLU	CD	-4.713	24.860	-0.108
	195	GLU	CD1	-3.118	24.960	0.185	195	GLU	CD2	-5.138	24.520	0.783
50	196	LEU	M	-0.829	25.264	-3.870	196	LEU	CA	8.241	25.929	-4.664
	196	LEU	C	0.228	25.376	-6.059	196	LEU	B	0.303	24.123	-6.153
	196	LEU	CB	1.340	25.739	-3.894	196	LEU	CG	2.770	26.178	-4.643
	196	LEU	CD1	2.739	27.716	-6.639	196	LEU	CD2	4.827	25.721	-3.911
	197	ASP	M	0.140	26.208	-7.893	197	ASP	CA	8.032	25.774	-8.488
55	197	ASP	C	1.387	25.738	-9.293	197	ASP	B	1.653	24.734	-9.916
	197	ASP	CB	-1.067	26.398	-9.191	197	ASP	CG	-2.486	26.351	-8.549
	197	ASP	CD1	-2.804	25.155	-8.334	197	ASP	CD2	-3.035	27.327	-8.888
	198	VAL	M	2.813	26.889	-9.344	198	VAL	CA	3.204	26.978	-10.109
	198	VAL	C	4.187	27.950	-9.514	198	VAL	B	3.752	28.699	-8.587
60	198	VAL	CB	2.894	27.476	-11.637	198	VAL	CG1	1.930	26.726	-12.537
	198	VAL	CG2	2.337	28.919	-11.684	199	MEY	M	5.374	27.916	-19.816
	199	MEY	CA	6.439	28.802	-9.498	199	MEY	C	6.849	28.810	-19.578
	199	MEY	B	6.696	29.518	-11.793	199	MEY	CB	7.660	27.970	-9.877
	199	MEY	CG	7.361	26.849	-8.139	199	MEY	CD	6.753	27.449	-6.568
65	199	MEY	CB	8.227	27.753	-8.137	200	ALA	M	7.426	30.942	-19.183
	200	ALA	CA	7.991	31.929	-11.033	200	ALA	C	9.888	32.666	-18.272
	200	ALA	B	8.127	32.924	-9.868	200	ALA	CB	6.932	32.870	-11.630

	201	PBC M	0.927	33.495	-10.951	201	PBC CA	11.013	34.130	-10.231
	201	PBC C	10.450	35.127	-9.231	201	PBC B	0.579	35.907	-9.682
	201	PBC CB	11.017	34.723	-11.400	201	PBC CC	11.392	34.040	-12.670
	201	PBC CD	0.941	33.616	-12.409	202	GLY M	10.920	35.204	-8.021
	202	GLY CA	10.473	34.234	-7.044	202	GLY C	11.300	34.610	-4.115
	202	GLY D	11.352	37.124	-4.979	203	VAL M	12.015	36.303	-6.613
5	203	VAL CA	13.940	36.929	-3.710	203	VAL C	14.706	38.017	-6.469
	203	VAL C	13.133	37.731	-7.503	203	VAL CE	14.014	35.600	-5.351
	203	VAL CG1	16.096	36.106	-4.612	203	VAL CG2	14.979	34.741	-4.378
	204	ILE M	14.963	39.102	-3.039	204	ILE CA	10.572	40.201	-6.407
	204	ILE C	11.047	40.619	-7.072	204	ILE C	15.706	40.605	-8.819
	204	ILE CB	17.087	39.974	-6.324	204	ILE CC	17.732	41.106	-6.472
	205	ILE M	13.771	40.045	-8.008	205	ILE CA	13.069	41.234	-9.235
	205	ILE C	13.207	42.749	-9.470	205	ILE D	12.675	43.490	-5.040
10	205	ILE CB	11.132	40.013	-9.144	205	ILE CG1	11.436	39.336	-8.810
	205	ILE CG2	10.899	41.211	-10.467	205	ILE CD1	12.257	38.412	-9.771
	206	GLN M	13.956	41.945	-10.409	206	GLN CA	14.204	44.517	-10.034
	206	GLN C	13.002	44.970	-11.630	206	GLN D	12.649	44.318	-12.621
	206	GLN CB	15.453	44.708	-11.740	206	GLN CC	16.604	44.143	-10.000
	206	GLN CD	17.205	43.145	-10.007	206	GLN CD1	10.320	44.936	-9.353
	206	GLN CD2	16.356	46.260	-9.037	207	SEP M	12.359	46.064	-11.214
	207	SEP CA	11.217	46.571	-11.907	207	SEP C	11.009	40.093	-11.749
15	207	SEP D	11.910	48.457	-11.004	207	SEP CB	9.910	43.033	-11.569
	207	SEP CC	8.993	46.034	-12.613	208	THR M	10.054	48.404	-12.326
	208	THR CG1	9.171	50.339	-14.734	208	THR CG2	7.570	49.414	-13.144
	208	THR CB	8.620	50.415	-13.357	208	THR CA	0.675	50.092	-12.173
	208	THR C	9.107	50.400	-10.003	208	THR D	8.423	49.037	-10.049
	209	LEU M	9.616	51.613	-10.220	209	LEU CA	9.192	52.150	-8.950
	209	LEU C	8.673	53.610	-9.262	209	LEU D	9.140	54.227	-10.222
	209	LEU CB	10.333	52.192	-7.938	209	LEU CC	10.004	50.016	-7.416
20	209	LEU CD1	13.968	51.114	-6.472	209	LEU CD2	9.607	50.202	-6.649
	210	PRO M	7.790	54.139	-8.444	210	PRO CA	7.273	55.517	-8.649
	210	PRO C	8.303	56.573	-8.639	210	PRO D	9.491	56.445	-8.104
	210	PRO CB	4.302	55.733	-7.917	210	PRO CC	0.004	54.379	-4.946
	210	PRO CD	7.193	53.491	-7.271	211	GLY M	0.077	57.665	-9.335
	211	GLY CA	9.069	50.763	-9.410	211	GLY C	10.094	50.434	-10.090
	211	GLY D	11.176	59.005	-10.209	212	ASN M	9.051	57.770	-11.907
	212	ASN CA	10.903	57.422	-12.643	212	ASN C	12.039	54.753	-12.056
25	212	ASN C	13.100	57.101	-12.420	212	ASN CB	11.224	50.395	-13.499
	212	ASN CC	13.003	50.105	-14.014	212	ASN CD1	11.053	57.054	-15.323
	212	ASN CD2	12.273	59.159	-10.376	213	LYS M	11.003	55.749	-11.247
	213	LYS CA	12.010	54.046	-10.337	213	LYS C	12.660	53.459	-10.066
	213	LYS D	11.775	53.039	-11.613	213	LYS CB	12.769	55.241	-9.059
	213	LYS CC	13.206	56.694	-8.767	213	LYS CD	13.146	57.030	-7.312
	213	LYS CE	14.105	50.210	-6.070	213	LYS CE	10.040	50.705	-7.921
	214	TYR M	13.601	52.703	-10.444	214	TYR CA	13.003	51.246	-10.722
30	214	TYR C	14.383	50.600	-9.009	214	TYR D	10.211	51.253	-8.017
	214	TYR CB	14.641	50.911	-11.904	214	TYR CC	14.130	51.621	-13.746
	214	TYR CD1	14.609	52.047	-13.678	214	TYR CD2	13.129	51.065	-14.014
	214	TYR CE1	14.230	53.475	-14.014	214	TYR CE2	12.654	51.669	-15.178
	214	TYR CE	13.204	52.095	-15.050	214	TYR DM	12.756	53.450	-16.696
	215	GLY M	14.930	49.347	-9.100	215	GLY CA	14.622	48.772	-7.005
	215	GLY C	14.130	47.325	-7.749	215	GLY D	13.249	46.917	-8.521
	216	ALA M	14.010	46.010	-6.031	216	ALA CA	14.654	45.203	-6.701
35	216	ALA C	13.602	44.922	-5.712	216	ALA D	13.940	45.527	-4.470
	216	ALA CB	15.715	44.354	-6.007	217	TYR M	12.700	43.022	-5.375
	217	TYR CA	13.964	43.600	-4.440	217	TYR C	12.033	43.920	-4.547
	217	TYR D	12.202	41.642	-5.016	217	TYR CE	10.473	43.062	-4.570
	217	TYR CC	10.137	40.291	-4.214	217	TYR CD1	10.046	45.991	-3.236
	217	TYR CD2	9.014	45.933	-4.705	217	TYR CE1	10.439	47.267	-2.700
	217	TYR CE2	8.654	47.210	-4.301	217	TYR CE	9.310	47.002	-3.391
	217	TYR DM	8.953	49.140	-2.900	218	ASN M	13.700	41.306	-3.391
40	218	ASN CA	11.040	39.042	-3.227	218	ASN C	10.204	39.636	-2.749

45

50

55

EP 0 251 446 B1

	218	ASL B	9.743	43.347	-1.017	218	ASL C0	12.953	30.340	-2.134
	218	ASL CG	14.831	39.566	-2.343	218	ASL DD1	14.612	39.789	-3.422
	218	ASL DD2	10.668	39.644	-2.163	218	SLY B	0.478	39.934	-2.289
	219	SLY CA	0.382	38.130	-2.649	219	SLY C	7.578	37.304	-3.681
5	219	SLY D	7.873	37.802	-4.876	220	THB B	6.501	36.438	-3.285
	220	THB CA	8.697	35.936	-4.179	220	THB C	4.879	37.044	-4.864
	220	THB C	6.417	36.742	-5.918	220	THB C0	4.825	34.819	-3.926
	220	THB DD1	6.136	38.543	-2.451	220	THB CG2	5.704	33.896	-3.980
	221	SLB B	4.738	38.238	-4.303	221	SLB CA	3.984	39.201	-5.369
	221	SLB C	4.760	39.641	-6.383	221	SLB D	4.117	40.208	-7.277
	221	SLB C0	3.323	40.383	-4.544	221	SLB C0	3.435	40.282	-3.149
	222	MEY B	6.063	39.389	-4.485	222	MEY C0	6.471	42.771	-9.173
10	222	MEY SD	7.768	41.933	-4.993	222	MEY C0	8.504	41.399	-4.402
	222	MEY C0	8.351	40.818	-7.218	222	MEY CA	6.916	39.670	-7.638
	222	MEY C	6.877	38.435	-8.567	222	MEY D	7.084	39.567	-9.775
	223	ALA B	8.554	37.244	-8.041	223	ALA CA	6.469	36.020	-8.885
	223	ALA C	8.200	36.068	-9.707	223	ALA D	8.133	35.948	-10.929
	223	ALA C0	6.909	34.807	-7.923	224	SLB B	4.076	36.360	-9.838
	224	SLB CA	2.788	36.689	-8.703	224	SLB C	2.641	37.161	-11.039
	224	SLB D	2.145	36.593	-12.037	224	SLB C0	1.801	36.995	-8.603
15	224	SLB D0	6.492	36.899	-9.137	225	PRC B	3.156	38.411	-11.159
	225	PRC CA	3.893	39.130	-12.439	225	PRC C	3.764	38.489	-13.424
	225	PRC D	3.406	38.650	-14.804	225	PRC C0	3.653	40.811	-12.054
	225	PRC CG	4.411	40.402	-10.764	225	PRC C0	3.733	39.124	-10.834
	226	MIS B	4.769	37.624	-13.299	226	MIS CA	5.446	36.879	-14.362
	226	MIS C	4.618	35.947	-15.061	226	MIS D	4.625	35.809	-16.293
	226	MIS C0	6.008	36.046	-13.765	226	MIS C0	7.814	36.859	-13.358
	226	MIS DD1	8.048	37.681	-12.170	226	MIS CG2	8.981	37.118	-14.167
20	226	MIS CG1	9.279	38.952	-12.236	226	MIS ME2	9.771	37.866	-13.463
	227	VAL B	3.593	35.366	-14.199	227	VAL CA	2.983	34.388	-14.727
	227	VAL C	1.479	33.197	-15.621	227	VAL D	3.018	34.773	-16.490
	227	VAL C0	2.103	33.644	-13.619	227	VAL CG1	3.076	32.476	-14.246
	227	VAL CG2	3.204	32.641	-12.891	228	ALA B	1.003	36.242	-14.814
	228	ALA CA	0.011	37.189	-15.517	228	ALA C	0.543	37.538	-16.868
	228	ALA D	-0.233	37.433	-17.828	228	ALA C0	-0.307	38.333	-14.668
	229	GLY B	1.791	38.028	-16.943	229	GLY CA	2.352	36.408	-18.239
25	229	GLY C	2.420	37.197	-19.187	229	GLY D	2.189	37.375	-20.384
	230	ALA B	2.711	35.981	-18.446	230	ALA CA	2.794	34.801	-19.944
	230	ALA C	1.424	34.500	-20.153	230	ALA D	1.380	34.203	-21.343
	230	ALA C0	3.298	33.624	-18.709	231	ALA B	0.385	34.623	-19.328
	231	ALA CA	-1.010	34.416	-19.744	231	ALA C	-1.256	35.423	-20.064
	231	ALA D	-1.908	35.856	-21.852	231	ALA C0	-1.032	34.664	-18.849
	232	ALA B	-0.778	36.457	-20.721	232	ALA CA	-1.013	37.663	-21.792
	232	ALA C	-0.281	37.284	-23.078	232	ALA D	-0.841	37.901	-24.187
30	232	ALA C0	-0.742	39.121	-21.377	233	LEU B	0.935	36.724	-22.967
	233	LEU CA	1.617	36.293	-24.209	233	LEU C	0.821	35.169	-24.880
	233	LEU D	0.696	35.231	-26.111	233	LEU C0	3.063	35.877	-23.907
	233	LEU CG	3.996	36.994	-23.653	233	LEU CG1	5.239	36.362	-22.921
	233	LEU CG2	4.241	37.853	-24.680	234	ILE B	0.357	34.199	-24.047
	234	ILE CG1	0.306	30.664	-21.637	234	ILE CG2	0.454	31.223	-23.189
	234	ILE C0	-0.811	32.014	-23.570	234	ILE C	-1.303	30.900	-24.891
	234	ILE CA	-0.406	32.076	-24.444	234	ILE C	-1.621	31.197	-25.434
35	234	ILE D	-1.883	33.144	-26.344	235	LEU B	-2.390	34.463	-24.779
	235	LEU CA	-3.396	35.028	-25.423	235	LEU C	-3.158	35.143	-26.672
	235	LEU D	-4.109	35.914	-27.589	235	LEU C0	-4.432	35.765	-24.378
	235	LEU CG	-5.140	36.899	-23.342	235	LEU CG1	-5.652	35.683	-22.165
	235	LEU CG2	-6.252	34.138	-24.120	236	SLB B	-2.094	34.638	-26.798
	236	SLB CA	-3.764	37.237	-27.986	236	SLB C	-3.491	36.292	-29.166
	236	SLB D	-3.746	36.634	-30.290	236	SLB C0	-0.653	34.234	-27.733
	236	SLB D0	0.599	37.571	-27.582	237	LYS B	-1.346	35.067	-28.881
40	237	LYS CA	-0.846	34.033	-29.952	237	LYS C	-2.113	33.177	-30.249
	237	LYS D	-2.378	32.951	-31.444	237	LYS C0	0.272	33.112	-29.551
	237	LYS CG	0.677	32.240	-30.716	237	LYS C0	2.020	31.925	-30.442

5	237	LV5	CE	2.345	20.762	-21.729	237	LV5	ME	2.925	29.848	-21.196
	238	M15	M	-2.931	21.989	-29.312	238	M15	CA	-4.168	22.163	-29.179
	239	M15	C	-3.334	22.099	-28.697	239	M15	O	-5.713	22.914	-27.562
	240	M15	CB	-3.948	20.862	-28.931	240	M15	CE	-3.089	29.921	-29.237
	241	M15	MC1	-1.707	29.679	-28.835	241	M15	CD2	-2.137	29.258	-29.394
	242	M15	CE1	-1.886	28.851	-29.642	242	M15	MF2	-1.948	28.688	-28.199
	243	PRD	M	-3.848	22.917	-29.385	243	PRD	CA	-6.988	24.779	-28.773
	244	PRD	C	-8.204	24.532	-28.332	244	PRD	O	-8.949	24.919	-27.662
	245	PRD	CB	-7.818	25.977	-29.713	245	PRD	CE	-6.666	23.294	-21.827
	246	PRD	CD	-3.436	26.439	-29.668	246	AS4	M	-8.386	22.969	-29.227
	247	AS4	CA	-9.529	22.041	-29.216	247	AS4	C	-9.589	21.189	-27.988
	248	AS4	M	-10.340	20.610	-27.576	248	AS4	CB	-9.493	21.249	-29.535
	249	AS4	CE	-7.971	20.827	-29.889	249	AS4	CD1	-7.898	21.990	-21.147
10	250	AS4	MD2	-7.675	29.309	-26.926	251	TRP	M	-8.354	21.806	-27.384
	251	TRP	CA	-8.304	20.124	-26.125	251	TRP	C	-9.186	20.638	-24.936
	252	TRP	O	-9.843	21.833	-26.686	252	TRP	CB	-6.879	29.835	-28.679
	253	TRP	CE	-6.894	28.953	-26.937	253	TRP	CD1	-6.338	28.433	-27.818
	254	TRP	CD2	-6.839	28.324	-26.183	254	TRP	MF1	-3.362	27.847	-28.211
	255	TRP	CE2	-6.414	27.476	-27.218	255	TRP	CE3	-4.897	28.408	-24.981
	256	TRP	CE2	-3.193	26.786	-27.174	256	TRP	CE3	-2.912	27.667	-24.943
	257	TRP	CE2	-2.479	26.873	-26.089	257	TMR	M	-8.727	29.781	-24.142
15	258	TMR	CA	-10.438	25.119	-22.911	258	TMR	C	-9.469	28.176	-21.747
	259	TMR	O	-8.333	29.674	-21.937	259	TMR	CB	-11.579	29.032	-22.675
	260	TMR	CD1	-10.837	27.786	-22.476	260	TMR	CE2	-12.494	28.987	-23.898
	261	AS4	M	-9.946	20.819	-28.611	261	AS4	MD2	-11.787	28.684	-18.747
	262	AS4	MD1	-11.465	21.813	-26.788	262	AS4	CE	-11.893	21.131	-17.988
	263	AS4	CB	-9.788	21.830	-28.332	263	AS4	CA	-9.893	20.731	-19.444
	264	AS4	C	-8.637	29.353	-29.018	264	AS4	O	-7.893	29.136	-18.448
	265	TMR	M	-9.564	28.162	-29.233	265	TMR	CA	-9.381	24.934	-19.859
20	266	TMR	C	-8.133	26.393	-29.882	266	TMR	O	-7.324	28.757	-19.111
	267	TMR	CB	-10.665	26.888	-29.494	267	TMR	CD1	-11.735	24.675	-18.684
	268	TMR	CE2	-10.553	24.995	-29.159	268	GLM	M	-8.582	24.716	-21.075
	269	GLM	CA	-6.964	26.362	-21.962	269	GLM	C	-3.647	27.820	-21.520
	270	GLM	O	-6.373	26.393	-21.447	270	GLM	CB	-7.330	24.899	-23.397
	271	GLM	CE	-8.165	25.521	-23.919	271	GLM	CD	-8.493	29.873	-25.428
	272	GLM	CE1	-9.386	24.769	-23.727	272	GLM	MD2	-7.745	23.312	-26.378
25	273	VAL	M	-5.697	28.304	-21.218	273	VAL	CA	-4.477	29.848	-29.778
	274	VAL	C	-2.936	28.462	-29.467	274	VAL	O	-2.789	28.227	-19.341
	275	VAL	CB	-4.779	20.555	-28.671	275	VAL	CD1	-3.544	21.272	-28.827
	276	VAL	CE2	-5.169	21.238	-21.959	276	ARG	M	-4.787	28.240	-18.462
	277	ARG	CA	-4.385	27.714	-17.168	277	ARG	C	-3.770	24.212	-17.348
	278	ARG	O	-2.708	25.985	-16.764	278	ARG	CB	-3.533	27.667	-16.149
	279	ARG	CE	-6.987	27.095	-14.892	279	ARG	CD	-6.856	27.179	-12.793
	280	ARG	ME	-5.440	26.757	-12.546	280	ARG	CE	-3.893	24.866	-12.315
	281	ARG	MD1	-7.864	27.484	-11.219	281	ARG	MD2	-3.177	28.428	-18.770
30	282	SER	M	-4.610	25.909	-18.131	282	SER	CA	-4.839	24.131	-18.426
	283	SER	C	-2.637	24.886	-18.872	283	SER	O	-1.848	23.253	-18.983
	284	SER	CB	-5.034	23.488	-19.372	284	SER	CD	-6.166	23.895	-18.832
	285	SER	M	-2.500	24.853	-20.136	285	SER	CB	-1.223	24.874	-28.811
	286	SER	C	-0.071	25.302	-19.940	286	SER	O	1.826	24.785	-28.049
	287	SER	CD	-1.369	29.758	-22.868	287	SER	CE	-9.389	25.419	-22.956
	288	LEU	M	-0.289	26.333	-19.160	288	LEU	CD2	1.824	29.814	-18.222
35	289	LEU	CD1	-0.373	20.453	-17.268	289	LEU	CE	0.352	29.438	-18.181
	290	LEU	CB	0.178	28.863	-17.903	290	LEU	CA	0.718	24.837	-18.218
	291	LEU	C	1.092	28.694	-17.263	291	LEU	C	2.293	25.421	-17.832
	292	GLM	M	0.968	25.857	-16.714	292	GLM	MD2	-2.750	25.312	-12.237
	293	GLM	CE1	-2.819	23.424	-12.931	293	GLM	CD	-2.368	24.850	-13.836
	294	GLM	CE	-1.218	24.814	-13.994	294	GLM	CB	-0.887	23.621	-14.877
	295	GLM	CA	0.581	23.941	-13.763	295	GLM	C	0.959	22.664	-18.361
	296	GLM	O	1.743	22.814	-13.616	296	AS4	M	0.633	22.394	-17.595
40	297	AS4	CA	1.892	21.204	-18.282	297	AS4	C	2.394	21.359	-18.993
	298	AS4	O	2.859	20.442	-19.768	298	AS4	CB	0.084	28.780	-19.282
	299	AS4	CE	-1.836	19.926	-18.573	299	AS4	CD1	-0.836	19.385	-17.582

45

50

55

5	212	AS4	DC1	-2.234	21.824	-19.141	253	YH4	h	3.818	22.581	-18.921
	253	YH4	CA	4.236	22.717	-19.713	253	YH4	z	9.361	23.247	-18.818
	253	YH4	D	6.346	23.733	-19.621	253	YH4	CB	4.884	23.671	-18.982
	253	YH4	DC1	3.593	24.937	-20.428	253	YH4	CC2	3.347	23.130	-22.032
	254	YH4	h	3.218	21.177	-17.911	254	YH4	CA	6.214	23.612	-16.988
10	254	YH4	C	7.466	27.700	-16.612	254	YH4	D	7.403	21.980	-17.893
	254	YH4	CB	5.664	23.938	-19.132	254	YH4	DC1	9.129	22.178	-18.040
	254	YH4	CC2	4.530	24.949	-14.802	255	YH4	h	8.499	23.296	-14.876
	255	YH4	CA	9.771	22.894	-19.817	255	YH4	C	9.621	22.031	-14.614
	255	YH4	D	9.439	22.784	-13.674	255	YH4	CB	11.980	23.453	-18.897
15	255	YH4	DC1	11.982	23.701	-17.921	255	YH4	CC2	12.288	22.628	-19.406
	256	LY3	h	9.604	20.702	-14.314	256	LY3	CA	9.364	20.063	-13.010
	256	LY3	C	10.920	20.333	-12.063	256	LY3	D	11.662	20.274	-12.592
	256	LY3	CB	9.074	18.990	-13.249	256	LY3	CC	9.818	17.803	-11.921
	256	LY3	CD	10.286	16.848	-11.777	256	LY3	CE	10.212	19.940	-18.623
20	256	LY3	CE	9.243	14.869	-11.054	257	LEU	h	10.212	20.674	-18.824
	257	LEU	CA	11.272	21.031	-9.893	257	LEU	C	11.210	20.232	-8.614
	257	LEU	D	12.096	20.865	-7.732	257	LEU	CB	11.187	22.947	-9.522
	257	LEU	CC	11.337	23.620	-10.968	257	LEU	CD1	11.248	25.083	-9.921
	257	LEU	CC2	12.678	23.468	-11.323	258	GLY	h	10.431	19.282	-8.298
25	258	GLY	CA	10.602	11.793	-6.879	258	GLY	C	9.168	18.703	-4.373
	258	GLY	D	8.283	18.936	-7.202	259	ASP	h	9.624	18.282	-5.150
	259	ASP	CA	7.737	17.896	-4.916	259	ASP	C	6.699	18.941	-4.789
	259	ASP	D	6.859	20.039	-4.214	259	ASP	CB	7.994	17.840	-3.033
	259	ASP	CC	6.781	17.128	-2.241	259	ASP	CD1	8.611	17.927	-2.314
30	259	ASP	CC2	7.898	16.391	-1.321	260	SEP	h	8.860	18.610	-5.312
	260	SEP	CA	6.681	19.587	-1.529	260	SEP	C	4.948	20.367	-4.299
	260	SEP	D	3.900	21.903	-4.446	260	SEP	CB	3.943	18.919	-6.289
	260	SEP	CC	2.743	17.937	-5.448	261	PHE	h	4.241	19.778	-3.112
	261	PHE	CA	3.831	28.481	-1.889	261	PHE	C	4.544	21.846	-1.863
35	261	PHE	D	3.944	22.848	-1.432	261	PHE	CB	4.053	19.749	-8.563
	261	PHE	CC	3.949	20.337	9.719	261	PHE	CD1	2.296	20.163	1.123
	261	PHE	CC2	4.401	21.840	1.398	261	PHE	CE1	1.737	28.717	2.319
	261	PHE	CE2	3.943	21.802	2.748	261	PHE	CE2	2.403	21.469	3.114
	262	TYR	h	9.778	21.798	-2.303	262	TYR	CA	6.688	22.914	-2.251
40	262	TYR	C	6.820	23.619	-3.549	262	TYR	D	7.201	24.853	-3.293
	262	TYR	CB	8.122	22.439	-1.851	262	TYR	CC	8.146	21.892	-8.454
	262	TYR	CD1	8.084	20.484	-0.364	262	TYR	CD2	8.149	22.649	0.498
	262	TYR	CE1	8.062	19.873	0.862	262	TYR	CE2	8.114	22.069	1.962
	262	TYR	CE2	8.069	20.672	2.018	262	TYR	DM	7.943	20.029	3.209
45	263	TYR	h	6.626	23.104	-4.493	263	TYR	CA	6.812	23.693	-4.822
	263	TYR	C	8.626	23.680	-6.956	263	TYR	D	9.781	24.117	-8.111
	263	TYR	CB	7.928	22.768	-6.681	263	TYR	CC	9.279	23.033	-6.068
	263	TYR	CD1	10.064	24.044	-6.637	263	TYR	CD2	9.800	22.342	-4.999
	263	TYR	CE1	11.339	24.328	-6.168	263	TYR	CE2	11.062	22.668	-4.491
50	263	TYR	CE2	11.838	23.618	-8.106	263	TYR	DM	12.863	23.949	-6.897
	264	GLY	h	4.671	23.161	-6.816	264	GLY	CA	3.301	23.064	-7.412
	264	GLY	C	3.847	22.196	-8.536	264	GLY	D	4.647	21.274	-8.363
	264	LY3	h	3.436	22.477	-9.734	264	LY3	CA	3.834	21.788	-10.971
	264	LY3	C	5.188	22.232	-11.464	264	LY3	D	3.684	21.863	-12.384
55	264	LY3	CB	2.789	22.071	-12.844	264	LY3	CC	1.490	21.963	-11.309
	264	LY3	CD	0.710	20.948	-12.079	264	LY3	CE	-0.692	20.496	-11.991
	264	LY3	CE2	-1.678	20.737	-12.489	266	GLY	h	3.787	23.224	-10.817
	266	GLY	CA	7.120	23.812	-11.323	266	GLY	C	7.193	29.882	-11.818
	266	GLY	D	6.177	23.793	-11.648	267	LEU	h	8.262	29.334	-12.480
60	267	LEU	CA	8.490	16.460	-13.097	267	LEU	C	7.804	26.771	-14.437
	267	LEU	D	7.953	23.909	-19.298	267	LEU	CB	19.810	26.893	-13.214
	267	LEU	CC	10.432	28.640	-14.858	267	LEU	CD1	10.896	29.331	-13.230
	267	LEU	CC2	11.924	27.921	-14.327	268	ILE	h	7.864	27.863	-16.632
	268	ILE	CA	4.404	18.038	-13.944	268	ILE	C	7.436	28.246	-17.863
65	268	ILE	D	8.539	28.793	-16.912	268	ILE	CB	8.969	29.210	-19.899
	268	ILE	CC1	6.099	30.541	-19.992	268	ILE	CC2	4.243	28.923	-14.867
	268	ILE	CD1	8.391	31.746	-16.262	269	ASN	h	7.887	27.843	-18.237

	269	Asn	CA	7.802	27.979	-21.459	269	Asn	C	26.879	28.954	-21.401
	269	Asn	D	1.945	27.162	-21.942	269	Asn	CB	26.492	26.413	-21.931
	269	Asn	CE	4.161	26.951	-21.216	269	Asn	CD1	8.993	27.624	-22.122
	269	Asn	CD2	11.011	25.796	-21.472	270	Val	H	6.998	28.388	-22.724
	270	Val	CA	8.313	27.418	-21.414	270	Val	E	6.891	28.987	-23.656
	270	Val	D	8.917	27.969	-22.072	270	Val	CB	8.646	28.719	-23.627
5	270	Val	CE1	6.849	22.717	-21.676	270	Val	CE2	6.420	22.542	-22.732
	271	Gln	H	7.323	29.761	-23.531	271	Gln	CA	7.892	29.278	-24.278
	271	Gln	C	6.869	27.714	-21.833	271	Gln	D	6.213	27.866	-24.091
	271	Gln	CB	8.104	25.220	-24.964	271	Gln	CE	9.486	28.018	-24.289
	271	Gln	CD	10.901	28.315	-26.082	271	Gln	CE1	11.261	28.549	-27.110
	271	Gln	CE2	11.702	26.913	-26.916	272	Ala	H	6.977	28.394	-24.002
	272	Ala	CA	6.274	25.712	-24.440	272	Ala	C	6.721	27.918	-24.241
	272	Ala	D	4.848	23.503	-23.101	272	Ala	CB	6.743	24.742	-23.172
10	272	Ala	H	4.207	24.061	-25.131	273	Ala	CA	2.949	24.381	-22.854
	273	Ala	C	4.041	27.671	-24.620	273	Ala	D	2.149	27.219	-24.181
	273	Ala	CB	4.716	27.773	-21.333	274	Ala	H	5.783	28.764	-24.742
	274	Ala	CB	3.932	30.341	-24.218	274	Ala	CA	2.189	29.144	-23.147
	274	Ala	C	4.730	23.367	-27.090	274	Ala	D	9.920	28.749	-27.621
	275	Gln	H	2.810	27.144	-27.214	275	Gln	CA	4.848	28.388	-28.627
	275	Gln	C	4.141	27.261	-27.777	275	Gln	D	6.749	27.817	-29.616
	275	Gln	CD	7.111	27.341	-30.790	275	Gln	CB	4.646	27.774	-28.520
15	275	Gln	CE	2.931	24.644	-27.467	275	Gln	CE	-3.823	23.936	-27.631
	275	Gln	CE1	-1.374	23.841	-26.720	275	Gln	CE2	-4.113	23.411	-24.938

The above structural studies together with the kinetic data presented herein and elsewhere (Philipp, M., et al. (1983) *Mol. Cell. Biochem.* 51, 5-32; Svendsen, I.B. (1976) *Carlsberg Res. Comm.* 41, 237-291; Markland, S.F. Id; Stauffe, D.C., et al. (1965) *J. Biol. Chem.* 244, 5333-5338) indicate that the subsites in the binding cleft of subtilisin are capable of interacting with substrate amino acid residues from P-4 to P-2'.

The most extensively studied of the above residues are Gly166, Gly169 and Ala152. These amino acids were identified as residues within the S-1 subsite. As seen in Fig. 3, which is a stereoview of the S-1 subsite, Gly166 and Gly169 occupy positions at the bottom of the S-1 subsite, whereas Ala152 occupies a position near the top of S-1, close to the catalytic Ser221.

All 19 amino acid substitutions of Gly166 and Gly169 have been made. As will be indicated in the examples which follow, the preferred replacement amino acids for Gly166 and/or Gly169 will depend on the specific amino acid occupying the P-1 position of a given substrate.

The only substitutions of Ala152 presently made and analyzed comprise the replacement of Ala152 with Gly and Ser. The results of these substitutions on P-1 specificity will be presented in the examples.

In addition to those residues specifically associated with specificity for the P-1 substrate amino acid, Tyr104 has been identified as being involved with P-4 specificity. Substitutions at Phe189 and Tyr217, however, are expected to respectively effect P-2' and P-1' specificity.

The catalytic activity of subtilisin has also been modified by single amino acid substitutions at Asn155. The catalytic triad of subtilisin is shown in Fig. 4. As can be seen, Ser221, His64 and Asp32 are positioned to facilitate nucleophilic attack by the serine hydroxylate on the carbonyl of the scissile peptide bond. Crystallographic studies of subtilisin (Robertus, et al. (1972) *Biochem.* 11, 4293-4303; Matthews, et al. (1975) *J. Biol. Chem.* 250, 7120-7126; Poulos, et al. (1976) *J. Biol. Chem.* 250, 1097-1103) show that two hydrogen bonds are formed with the oxyanion of the substrate transition state. One hydrogen bond donor is from the catalytic serine-221 main-chain amide while the other is from one of the NE2 protons of the asparagine-155 side chain. See Fig. 4.

Asn155 was substituted with Ala, Asp, His, Glu and Thr. These substitutions were made to investigate the the stabilization of the charged tetrahedral intermediate of the transition state complex by the potential hydrogen bond between the side chain of Asn155 and the oxyanion of the intermediate. These particular substitutions caused large decreases in substrate turnover, *k_{cat}* (200 to 4,000 fold), marginal decreases in substrate binding *K_m* (up to 7 fold), and a loss in transition state stabilization energy of 2.2 to 4.7 kcal/mol. The retention of *K_m* and the drop in *k_{cat}* will make these mutant enzymes useful as binding proteins for specific; peptide sequences, the nature of which will be determined by the specificity of the precursor protease.

Various other amino acid residues have been identified which affect alkaline stability. In some cases, mutants having altered alkaline stability also have altered thermal stability.

In *B. amyloliquefaciens* subtilisin residues Asp36, Ile107, Lys170, Ser204 and Lys213 have been identified as residues which upon substitution with a different amino acid alter the alkaline stability of the mutated enzyme as compared to the precursor enzyme. The substitution of Asp36 with Ala and the substitution of Lys170 with Glu each resulted in a mutant enzyme having a lower alkaline stability as compared to the wild type subtilisin. When Ile107 was substituted with Val, Ser204 substituted with Cys, Arg or Leu or Lys213 substituted with Arg, the mutant subtilisin had a greater alkaline stability as compared

to the wild type subtilisin. However, the mutant Ser204P demonstrated a decrease in alkaline stability.

In addition, other residues, identified as being associated with the modification of other properties of subtilisin, also affect alkaline stability. These residues include Ser24, Met50, Glu156, Gly166, Gly169 and Tyr217. Specifically the following particular substitutions result in an increased alkaline stability: Ser24C, Met50F, Gly156Q or S, Gly166A, H, K, N or Q, Gly169S or A, and Tyr217F, K, R or L. The mutant Met50V, on the other hand, results in a decrease in the alkaline stability of the mutant subtilisin as compared to wild type subtilisin.

Other residues involved in alkaline stability based on the alkaline stability screen include Asp197 and Met222. Particular mutants include Asp197(R or A) and Met 222 (all other amino acids).

Various other residues have been identified as being involved in thermal stability as determined by the thermal stability screen herein. These residues include the above identified residues which effect alkaline stability and Met199 and Tyr21. These latter two residues are also believed to be important for alkaline stability. Mutants at these residues include I199 and F21.

The amino acid sequence of B. amyloliquefaciens subtilisin has also been modified by substituting two or more amino acids of the wild-type sequence. Six categories of multiply substituted mutant subtilisin have been identified. The first two categories comprise thermally and oxidatively stable mutants. The next three other categories comprise mutants which combine the useful properties of any of several single mutations of B. amyloliquefaciens subtilisin. The last category comprises mutants which have modified alkaline and/or thermal stability.

The first category comprises double mutants in which two cysteine residues have been substituted at various amino acid residue positions within the subtilisin molecule. Formation of disulfide bridges between the two substituted cysteine residues results in mutant subtilisins with altered thermal stability and catalytic activity. These mutants include A21/C22/C87 and C24/C87 which will be described in more detail in Example 11.

The second category of multiple subtilisin mutants comprises mutants which are stable in the presence of various oxidizing agents such as hydrogen peroxide or peracids. Examples 1 and 2 describe these mutants which include F50/I124/Q222, F50/I124, F50/Q222, F50/L124/Q222, I124/Q222 and L124/Q222.

The third category of multiple subtilisin mutants comprises mutants with substitutions at position 222 combined with various substitutions at positions 166 or 169. These mutants, for example, combine the property of oxidative stability of the A222 mutation with the altered substrate specificity of the various 166 or 169 substitutions. Such multiple mutants include A166/A222, A166/C222, F166/C222, K166/A222, K166/C222, V166/A222 and V166/C222. The K166/A222 mutant subtilisin, for example, has a kcat/Km ratio which is approximately two times greater than that of the single A222 mutant subtilisin when compared using a substrate with phenylalanine as the P-1 amino acid. This category of multiple mutant is described in more detail in Example 12.

The fourth category of multiple mutants combines substitutions at position 156 (Glu to Q or S) with the substitution of Lys at position 166. Either of these single mutations improve enzyme performance upon substrates with glutamate as the P-1, amino acid. When these single mutations are combined, the resulting multiple enzyme mutants perform better than either precursor. See Example 9.

The fifth category of multiple mutants contain the substitution of up to four amino acids of the B. amyloliquefaciens subtilisin sequence. These mutants have specific properties which are virtually identical to the properties of the subtilisin from B. licheniformis. The subtilisin from B. licheniformis differs from B. amyloliquefaciens subtilisin at 87 out of 275 amino acids. The multiple mutant F50/S156/A169/L217 was found to have similar substrate specificity and kinetics to the licheniformis enzyme. (See Example 13.) However, this is probably due to only three of the mutations (S156, A169 and L217) which are present in the substrate binding region of the enzyme. It is quite surprising that, by making only three changes out of the 87 different amino acids between the sequence of the two enzymes, the B. amyloliquifaciens enzyme was converted into an enzyme with properties similar to B. licheniformis enzyme. Other enzymes in this series include F50/Q156/N166/L217 and F50/S156/L217.

The sixth category of multiple mutants includes the combination of substitutions at position 107 (Ile to V) with the substitution of Lys at position 213 with Arg, and the combination of substitutions of position 204 (preferably Ser to C or L but also to all other amino acids) with the substitution of Lys at position 213 with R. Other multiple mutants which have altered alkaline stability include Q156/K166, Q156/N166, S156/K166, S156/N166 (previously identified as having altered substrate specificity), and F50/S156/A169/L217 (previously identified as a mutant of B. amyloliquifaciens subtilisin having properties similar to subtilisin from B. licheniformis). The mutant F50/V107/R213 was constructed based on the observed increase in alkaline stability for the single mutants F50, V107 and R213. It was determined that the V107/R213 mutant had an increased alkaline stability as compared to the wild type subtilisin. In this particular mutant, the increased

alkaline stability was the result of the cumulative stability of each of the individual mutations. Similarly, the mutant F50/V107/R213 had an even greater alkaline stability as compared to the V107/R213 mutant indicating that the increase in the alkaline stability due to the F50 mutation was also cumulative.

Table IV summarizes the multiple mutants which have been made including those not mentioned above.
 5 In addition, based in part on the above results, substitution at the following residues in subtilisin is expected to produce a multiple mutant having increased thermal and alkaline stability: Ser24, Met50, Ile107, Glu156, Gly166, Gly169, Ser204, Lys213, Gly215, and Tyr217.

TABLE IV

10

15

20

25

30

35

40

Double Mutants	Triple, Quadruple or Other Multiple
C22/C87	F50/I124/Q222
C24/C87	F50/L124/Q222
V45/V48	F50/L124/A222
C49/C94	A21/C22/C87
C49/C95	F50/S156/N166/L217
C50/C95	F50/Q156/N166/L217
C50/C110	F50/S156/A169/L217
F50/I124	F50/S156/L217
F50/Q222	F50/Q156/K166/L217
I124/Q222	F50/S156/K166/L217
Q156/D166	F50/Q156/K166/K217
Q156/K166	F50/S156/K166/K217
Q156/N166	F50/V107/R213
S156/D166	[S153/S156/A158/G159/S160/ Δ 161-164/I165/S166/A169/R170]
S156/K166	
S156/N166	L204/R213
S156/A169	R213/204A, E, Q, D, N, G, K, V, R, T, P, I, M, F, Y, W or H
A166/A222	
A166/C222	
F166/A222	V107/R213
F166/C222	
K166/A222	
K166/C222	
V166/A222	
V166/C222	
A169/A222	
A169/A222	
A169/C222	
A21/C22	

45

In addition to the above identified amino acid residues, other amino acid residues of subtilisin are also considered to be important with regard to substrate specificity. Mutation of each of these residues is expected to produce changes in the substrate specificity of subtilisin. Moreover, multiple mutations among these residues and among the previously identified residues are also expected to produce subtilisin mutants having novel substrate specificity.

50

Particularly important residues are His67, Ile107, Leu126 and Leu135. Mutation of His67 should alter the S-1' subsite, thereby altering the specificity of the mutant for the P-1' substrate residue. Changes at this position could also affect the pH activity profile of the mutant. This residue was identified based on the inventor's substrate modeling from product inhibitor complexes.

55

Ile107 is involved in P-4 binding. Mutation at this position thus should alter specificity for the P-4 substrate residue in addition to the observed effect on alkaline stability. Ile107 was also identified by molecular modeling from product inhibitor complexes.

The S-2 binding site includes the Leu126 residue. Modification at this position should therefore affect P-2 specificity. Moreover, this residue is believed to be important to convert subtilisin to an amino peptidase.

The pH activity profile should also be modified by appropriate substitution. These residues were identified from inspection of the refined model, the three dimensional structure from modeling studies. A longer side chain is expected to preclude binding of any side chain at the S-2 subsite. Therefore, binding would be restricted to subsites S-1, S-1', S-2', S-3' and cleavage would be forced to occur after the amino terminal peptide.

Leu135 is in the S-4 subsite and if mutated should alter substrate specificity for P-4 if mutated. This residue was identified by inspection of the three-dimensional structure and modeling based on the product inhibitor complex of F222.

In addition to these sites, specific amino acid residues within the segments 97-103, 126-129 and 213-215 are also believed to be important to substrate binding.

Segments 97-103 and 126-129 form an antiparallel beta sheet with the main chain of substrate residues P-4 through P-2. Mutating residues in those regions should affect the substrate orientation through main chain (enzyme) - main chain (substrate) interactions, since the main chain of these substrate residues do not interact with these particular residues within the S-4 through S-2 subsites.

Within the segment 97-103, Gly97 and Asp99 may be mutated to alter the position of residues 101-103 within the segment. Changes at these sites must be compatible, however. In *B. amyloliquifaciens* subtilisin Asp99 stabilizes a turn in the main chain tertiary folding that affects the direction of residues 101-103. *B. licheniformis* subtilisin Asp97, functions in an analogous manner.

In addition to Gly97 and Asp99, Ser101 interacts with Asp99 in *B. amyloliquifaciens* subtilisin to stabilize the same main chain turn. Alterations at this residue should alter the 101-103 main chain direction. Mutations at Glu103 are also expected to affect the 101-103 main chain direction.

The side chain of Gly102 interacts with the substrate P-3 amino acid. Side chains of substituted amino acids thus are expected to significantly affect specificity for the P-3 substrate amino acids.

All the amino acids within the 127-129 segment are considered important to substrate specificity. Gly127 is positioned such that its side chain interacts with the S-1 and S-3 subsites. Altering this residue thus should alter the specificity for P-1 and P-3 residues of the substrate.

The side chain of Gly128 comprises a part of both the S-2 and S-4 subsites. Altered specificity for P-2 and P-4 therefore would be expected upon mutation. Moreover, such mutation may convert subtilisin into an amino peptidase for the same reasons substitutions of Leu126 would be expected to produce that result.

The Pro129 residue is likely to restrict the conformational freedom of the sequence 126-133, residues which may play a major role in determining P-1 specificity. Replacing Pro may introduce more flexibility thereby broadening the range of binding capabilities of such mutants.

The side chain of Lys213 is located within the S-3 subsite. All of the amino acids within the 213-215 segment are also considered to be important to substrate specificity. Accordingly, altered P-3 substrate specificity is expected upon mutation of this residue.

The Tyr214 residue does not interact with substrate but is positioned such that it could affect the conformation of the hair pin loop 204-217.

Finally, mutation of the Gly215 residue should affect the S-3' subsite, and thereby alter P-3' specificity.

In addition to the above substitutions of amino acids, the insertion or deletion of one or more amino acids within the external loop comprising residues 152-172 may also affect specificity. This is because these residues may play a role in the "secondary contact region" described in the model of streptomyces subtilisin inhibitor complexed with subtilisin. Hirono, et al. (1984) *J. Mol. Biol.* 178, 389-413. Thermitase K has a deletion in this region, which eliminates several of these "secondary contact" residues. In particular, deletion of residues 161 through 164 is expected to produce a mutant subtilisin having modified substrate specificity. In addition, a rearrangement in this area induced by the deletion should alter the position of many residues involved in substrate binding, predominantly at P-1. This, in turn, should affect overall activity against proteinaceous substrates

The effect of deletion of residues 161 through 164 has been shown by comparing the activity of the wild type (WT) enzyme with a mutant enzyme containing this deletion as well as multiple substitutions (i.e., S153/S156/A158/G159/S160/ Δ 161-164/I165/S166/A169/R170). This produced the following results:

TABLE V

	kcat	Km	kcat/Km
WT	50	1.4×10^{-4}	3.6×10^5
Deletion mutant	8	5.0×10^{-6}	1.6×10^6

The WT has a k_{cat} 6 times greater than the deletion mutant but substrate binding is 28 fold tighter by the deletion mutant. The overall efficiency of the deletion mutant is thus 4.4 times higher than the WT enzyme.

All of these above identified residues which have yet to be substituted, deleted or inserted into are presented in Table VI.

TABLE VI

Substitution/Insertion/Deletion	
Residues	
His67	Ala152
Leu126	Ala153
Leu135	Gly154
Gly97	Asn155
Asp99	Gly156
Ser101	Gly157
Gly102	Gly160
Glu103	Thr158
Leu126	Ser159
Gly127	Ser161
Gly128	Ser162
Pro129	Ser163
Tyr214	Thr164
Gly215	Val165
Gly166	Gly169
Tyr167	Lys170
Pro168	Tyr171
	Pro172

The following disclosure is intended to serve as a representation of embodiments herein, and should not be construed as limiting the scope of this application. These specific examples disclose the construction of certain of the above identified mutants. The construction of the other mutants, however, is apparent from the disclosure herein and that presented in EPO Publication No. 0130756.

All literature citations are expressly incorporated by reference.

EXAMPLE 1

Identification of Peracid Oxidizable Residues of Subtilisin Q222 and L222

As shown in Figures 6A and 6B, organic peracid oxidants inactivate the mutant subtilisins Met222L and Met222Q (L222 and Q222). This example describes the identification of peracid oxidizable sites in these mutant subtilisins.

First, the type of amino acid involved in peracid oxidation was determined. Except under drastic conditions (Means, G.E., et al. (1971) Chemical Modifications of Proteins, Holden-Day, S.F., CA, pp. 160-162), organic peracids modify only methionine and tryptophan in subtilisin. Difference spectra of the enzyme over the 250nm to 350nm range were determined during an inactivation titration employing the reagent, diperdodecanoic acid (DPDA) as oxidant. Despite quantitative inactivation of the enzyme, no change in absorbance over this wavelength range was noted as shown in Figures 7A and 7B indicating that tryptophan was not oxidized. Fontana, A., et al. (1980) Methods in Peptide and Protein Sequence Analysis - (C. Birr ed.) Elsevier, New York, p. 309. The absence of tryptophan modification implied oxidation of one or more of the remaining methionines of *B. amyloliquefaciens* subtilisin. See Figure 1.

To confirm this result the recombinant subtilisin Met222F was cleaved with cyanogen bromide (CNBr) both before and after oxidation by DPDA. The peptides produced by CNBr cleavage were analyzed on high resolution SDS-pyridine peptide gels (SPG).

Subtilisin Met222F (F222) was oxidized in the following manner. Purified F222 was resuspended in 0.1 M sodium borate pH 9.5 at 10 mg/ml and was added to a final concentration of 26 diperdodecanoic acid

(DPDA) at 26 mg/ml was added to produce an effective active oxygen concentration of 30 ppm. The sample was incubated for at least 30 minutes at room temperature and then quenched with 0.1 volume of 1 M Tris pH 8.6 buffer to produce a final concentration of 0.1 M Tris pH 8.6). 3mM phenylmethylsulfonyl fluoride (PMSF) was added and 2.5 ml of the sample was applied to a Pharmacia PD10 column equilibrated in 10 mM sodium phosphate pH 6.2, 1 mM PMSF. 3.5 ml of 10 mM sodium phosphate pH6.2, 1mM PMSF was applied and the eluant collected.

F222 and DPDA oxidized F222 were precipitated with 9 volumes of acetone at -20°C. The samples were resuspended at 10 mg/ml in 8M urea in 88% formic acid and allowed to sit for 5 minutes. An equal volume of 200 mg/ml CNBr in 88% formic acid was added (5 mg/ml protein) and the samples incubated for 2 hours at room temperature in the dark. Prior to gel electrophoresis, the samples were lyophilized and resuspended at 2-5 mg/ml in sample buffer (1% pyridine, 5% NaDodSO₄, 5% glycerol and bromophenol blue) and disassociated at 95°C for 3 minutes.

The samples were electrophoresed on discontinuous polyacrylamide gels (Kyte, J., et al. (1953) Anal. Bioch. 133, 515-522). The gels were stained using the Pharmacia silver staining technique (Sammons, D.W., et al. (1981) Electrophoresis 2 135-141).

The results of this experiment are shown in Figure 8. As can be seen, F222 treated with CNBr only gives nine resolved bands on SPG. However, when F222 is also treated with DPDA prior to cleavage, bands X, 7 and 9 disappear whereas bands 5 and 6 are greatly increased in intensity.

In order to determine which of the methionines were effected, each of the CNBr peptides was isolated by reversed phase HPLC and further characterized. The buffer system in both Solvent A (aqueous) and Solvent B (organic) for all HPLC separations was 0.05% triethylamine/trifluoroacetic acid (TEA-TFA). In all cases unless noted, solvent A consisted of 0.05% TEA-TFA in H₂O, solvent B was 0.05% TEA-TFA in 1-propanol, and the flow rate was 0.5 ml/minute.

For HPLC analysis, two injections of 1 mg enzyme digest were used. Three samples were acetone precipitated, washed and dried. The dried 1 mg samples were resuspended at 10 mg/ml in 8M urea, 88% formic acid; an equal volume of 200 mg/ml CNBr in 88% formic acid was added (5 mg/ml protein). After incubation for 2 hours in the dark at room temperature, the samples were desalted on a 0.8 cm X 7 cm column of Tris Acryl GF05 coarse resin (IBF, Paris, France) equilibrated with 40% solvent B, 60% solvent A. 200 ul samples were applied at a flow rate of 1 ml a minute and 1.0-1.2 ml collected by monitoring the absorbance at 280nm. Prior to injection on the HPLC, each desalted sample was diluted with 3 volumes of solvent A. The samples were injected at 1.0 ml/min (2 minutes) and the flow then adjusted to 0.5 ml/min (100% A). After 2 minutes, a linear gradient to 60% B at 1.0% B/min was initiated. From each 1 mg run, the pooled peaks were sampled (50ul) and analyzed by gel electrophoresis as described above.

Each polypeptide isolated by reversed phase HPLC was further analyzed for homogeneity by SPG. The position of each peptide on the known gene sequence (Wells, J.A., et al. (1983) Nucleic Acids Res. 11 7911-7924) was obtained through a combination of amino acid compositional analysis and, where needed, amino terminal sequencing.

Prior to such analysis the following peptides were to rechromatographed.

1. CNBr peptides from F222 not treated with DPDA:

Peptide 5 was subjected to two additional reversed phase separations. The 10 cm C4 column was equilibrated to 80%A/ 20%B and the pooled sample applied and washed for 2 minutes. Next an 0.5% ml B/min gradient was initiated. Fractions from this separation were again rerun, this time on the 25 cm C4 column, and employing 0.05% TEA-TFA in acetonitrile/1-propanol (1:1) for solvent B. The gradient was identical to the one just described.

Peptide "X" was subjected to one additional separation after the initial chromatography. The sample was applied and washed for 2 minutes at 0.5ml/min (100%A), and a 0.5% ml B/min gradient was initiated.

Peptides 7 and 9 were rechromatographed in a similar manner to the first rerun of peptide 5.

Peptide 8 was purified to homogeneity after the initial separation.

2. CNBr Peptides from DPDA Oxidized F222:

Peptides 5 and 6 from a CNBr digest of the oxidized F222 were purified in the same manner as peptide 5 from the untreated enzyme.

Amino acid compositional analysis was obtained as follows. Samples (-1nM each amino acid) were dried, hydrolyzed in vacuo with 100 ul 6N HCl at 106°C for 24 hours and then dried in a Speed Vac. The samples were analyzed on a Beckmann 6300 AA analyzer employing ninhydrin detection.

Amino terminal sequence data was obtained as previously described (Rodriguez, H., et al. (1984) Anal. Biochem. 134, 538-547).

The results are shown in Table VII and Figure 9.

TABLE VII

Amino and COOH termini of CNBr fragments Terminus and Method		
Fragment	amino, method	COOH, method
X	1, sequence	50, composition
9	51, sequence	119, composition
7	125, sequence	199, composition
8	200, sequence	275, composition
5ox	1, sequence	119, composition
6ox	120, composition	199, composition

Peptides 5ox and 6ox refer to peptides 5 and 6 isolated from CNBr digests of the oxidized protein where their respective levels are enhanced.

From the data in Table VII and the comparison of SPG tracks for the oxidized and native protein digests in Figure 8, it is apparent that (1) Met50 is oxidized leading to the loss of peptides X and 9 and the appearance of 5; and (2) Met124 is also oxidized leading to the loss of peptide 7 and the accumulation of peptide 6. Thus oxidation of B. amyloliquifaciens subtilisin with the peracid, diperdocecanoic acid leads to the specific oxidation of methionine at residues 50 and 124.

EXAMPLE 2

Substitution at Met50 and Met124 in Subtilisin Met222Q

The choice of amino acid for substitution at Met50 was based on the available sequence data for subtilisins from B. licheniformis (Smith, E.C., et al. (1968) J. Biol. Chem. 243, 2184-2191), B.DY (Nedkov, P., et al. (1983) Hoppe Saylor's Z. Physiol. Chem. 364 1537-1540), B. amylosacchariticus (Markland, F.S., et al. (1967) J. Biol. Chem. 242 5198-5211) and B. subtilis (Stahl, M.L., et al. (1984) J. Bacteriol. 158, 411-418). In all cases, position 50 is a phenylalanine. See Figure 5. Therefore, Phe50 was chosen for construction.

At position 124, all known subtilisins possess a methionine. See Figure 5. Molecular modelling of the x-ray derived protein structure was therefore required to determine the most probable candidates for substitution. From all 19 candidates, isoleucine and leucine were chosen as the best residues to employ. In order to test whether or not modification at one site but not both was sufficient to increase oxidative stability, all possible combinations were built on the Q222 backbone (F50/Q222, I124/Q222, F50/I124/Q222).

A. Construction of Mutations Between Codons 45 and 50

All manipulations for cassette mutagenesis were carried out on pS4.5 using methods disclosed in EPO Publication No. 0130756 and Wells, J.A., et al. (1985) Gene 34, 315-323. The pΔ50 in Fig. 10, line 4, mutations was produced using the mutagenesis primer shown in Fig. 10, line 6, and employed an approach designated as restriction-purification which is described below. Briefly, a M13 template containing the subtilisin gene, M13mp11-SUBT was used for heteroduplex synthesis (Adelman, et al (1983), DNA 2, 183-193). Following transfection of JM101 (ATCC 33876), the 1.5 kb EcoRI-BamHI fragment containing the subtilisin gene was subcloned from M13mp11 SUBT rf into a recipient vector fragment of pBS42 the construction of which is described in EPO Publication No. 0130756. To enrich for the mutant sequence (pΔ50, line 4), the resulting plasmid pool was digested with KpnI, and linear molecules were purified by polyacrylamide gel electrophoresis. Linear molecules were ligated back to a circular form, and transformed into E. coli MM294 cells (ATCC 31446). Isolated plasmids were screened by restriction analysis for the KpnI site. KpnI⁺ plasmids were sequenced and confirmed the pΔ50 sequence. Asterisks in Figure 11 indicate the bases that are mutated from the wild type sequence (line 4). pΔ50 (line 4) was cut with StuI and EcoRI and the 0.5 Kb fragment containing the 5' half of the subtilisin gene was purified (fragment 1). pΔ50 (line 4) was digested with KpnI and EcoRI and the 4.0 Kb fragment containing the 3' half of the subtilisin gene and vector sequences was purified (fragment 2). Fragments 1 and 2 (line 5), and duplex DNA

cassettes coding for mutations desired (shaded sequence, line 6) were mixed in a molar ratio of 1:1:10, respectively. For the particular construction of this example the DNA cassette contained the triplet TTT for codon 50 which encodes Phe. This plasmid was designated pF50. The mutant subtilisin was designated F50.

5

B. Construction of Mutation Between Codons 122 and 127

The procedure of Example 2A was followed in substantial detail except that the mutagenesis primer of Figure 11, line 7 was used and restriction-purification for the EcoRV site in p Δ 124 was used. In addition, the DNA cassette (shaded sequence, Figure 11, line 6) contained the triplet ATT for codon 124 which encodes Ile and CTT for Leu. Those plasmids which contained the substitution of Ile for Met124 were designated pI124. The mutant subtilisin was designated I124.

15

C. Construction of Various F50/I124/Q222 Multiple Mutants

The triple mutant, F50/I124/Q222, was constructed from a three-way ligation in which each fragment contained one of the three mutations. The single mutant Q222 (pQ222) was prepared by cassette mutagenesis as described in EPO Publication No. 0130756. The F50 mutation was contained on a 2.2kb Avall to PvuII fragment from pF50; the I124 mutation was contained on a 260 bp PvuII to Avall fragment from pI124; and the Q222 mutation was contained on 2.7 kb Avall to Avall fragment from pQ222. The three fragments were ligated together and transformed into E. coli MM294 cells. Restriction analysis of plasmids from isolated transformants confirmed the construction. To analyze the final construction it was convenient that the Avall site at position 798 in the wild-type subtilisin gene was eliminated by the I124 construction.

The F50/Q222 and I124/Q222 mutants were constructed in a similar manner except that the appropriate fragment from pS4.5 was used for the final construction.

D. Oxidative Stability of Q222 Mutants

The above mutants were analyzed for stability to peracid oxidation. As shown in Fig. 12, upon incubation with diperoxidodecanoic acid (protein 2mg/mL, oxidant 75ppm[O]), both the I124/Q222 and the F50/I124/Q222 are completely stable whereas the F50/Q222 and the Q222 are inactivated. This indicates that conversion of Met124 to I124 in subtilisin Q222 is sufficient to confer resistance to organic peracid oxidants.

EXAMPLE 3

Subtilisin Mutants Having Altered Substrate Specificity-Hydrophobic Substitutions at Residues 166

Subtilisin contains an extended binding cleft which is hydrophobic in character. A conserved glycine at residue 166 was replaced with twelve non-ionic amino acids which can project their side-chains into the S-1 subsite. These mutants were constructed to determine the effect of changes in size and hydrophobicity on the binding of various substrates.

A. Kinetics for Hydrolysis of Substrates Having Altered P-1 Amino Acids by Subtilisin from B. Amyloliquefaciens

Wild-type subtilisin was purified from B. subtilis culture supernatants expressing the B. amyloliquefaciens subtilisin gene (Wells, J.A., et al. (1983) Nucleic Acids Res. 11, 7911-7925) as previously described (Estell, D.A., et al. (1985) J. Biol. Chem. 260, 6518-6521). Details of the synthesis of tetrapeptide substrates having the form succinyl-L-AlaL-AlaL-ProL-[X]-p-nitroanilide (where X is the P1 amino acid) are described by DelMar, E.G., et al. (1979) Anal. Biochem. 99, 316-320. Kinetic parameters, K_m (M) and k_{cat} (s⁻¹) were measured using a modified progress curve analysis (Estell, D.A., et al. (1985) J. Biol. Chem. 260, 6518-6521). Briefly, plots of rate versus product concentration were fit to the differential form of the rate equation using a non-linear regression algorithm. Errors in k_{cat} and K_m for all values reported are less than five percent. The various substrates in Table VIII are ranged in order of decreasing hydrophobicity. Nozaki, Y. (1971), J. Biol. Chem. 246, 2211-2217; Tanford C. (1978) Science 200, 1012).

TABLE VIII

P1 substrate Amino Acid	kcat(S ⁻¹)	1/Km(M ⁻¹)	kcat/Km (s ⁻¹ M ⁻¹)
Phe	50	7,100	360,000
Tyr	28	40,000	1,100,000
Leu	24	3,100	75,000
Met	13	9,400	120,000
His	7.9	1,600	13,000
Ala	1.9	5,500	11,000
Gly	0.003	8,300	21
Gln	3.2	2,200	7,100
Ser	2.8	1,500	4,200
Glu	0.54	32	16

The ratio of kcat/Km (also referred to as catalytic efficiency) is the apparent second order rate constant for the conversion of free enzyme plus substrate (E + S) to enzyme plus products (E + P) (Jencks, W.P., Catalysis in Chemistry and Enzymology (McGraw-Hill, 1969) pp. 321-436; Fersht, A., Enzyme Structure and Mechanism (Freeman, San Francisco, 1977) pp. 226-287). The log (kcat/Km) is proportional to transition state binding energy, ΔG^\ddagger . A plot of the log kcat/Km versus the hydrophobicity of the P1 side-chain (Figure 14) shows a strong correlation ($r = 0.98$), with the exception of the glycine substrate which shows evidence for non-productive binding. These data show that relative differences between transition-state binding energies can be accounted for by differences in P-1 side-chain hydrophobicity. When the transition-state binding energies are calculated for these substrates and plotted versus their respective side-chain hydrophobicities, the line slope is 1.2 (not shown). A slope greater than unity, as is also the case for chymotrypsin (Fersht, A., Enzyme Structure and Mechanism (Freeman, San Francisco, 1977) pp. 226-287; Harper, J.W., et al. (1984) Biochemistry, 23, 2995-3002), suggests that the P1 binding cleft is more hydrophobic than ethanol or dioxane solvents that were used to empirically determine the hydrophobicity of amino acids (Nozaki, Y., et al. J. Biol. Chem. (1971) 246, 2211-2217; Tanford, C. (1978) Science 200, 1012).

For amide hydrolysis by subtilisin, kcat can be interpreted as the acylation rate constant and Km as the dissociation constant, for the Michaelis complex (E•S), Ks. Gutfreund, H., et al (1956) Biochem. J. 63, 656. The fact that the log kcat, as well as log 1/Km, correlates with substrate hydrophobicity is consistent with proposals (Robertus, J.D., et al. (1972) Biochemistry 11, 2439-2449; Robertus, J.D., et al. (1972) Biochemistry 11, 4293-4303) that during the acylation step the P-1 side-chain moves deeper into the hydrophobic cleft as the substrate advances from the Michaelis complex (E•S) to the tetrahedral transition-state complex (E•S^{*}). However, these data can also be interpreted as the hydrophobicity of the P1 side-chain effecting the orientation, and thus the susceptibility of the scissile peptide bond to nucleophilic attack by the hydroxyl group of the catalytic Ser221.

The dependence of kcat/Km on P-1 side chain hydrophobicity suggested that the kcat/Km for hydrophobic substrates may be increased by increasing the hydrophobicity of the S-1 binding subsite. To test this hypothesis, hydrophobic amino acid substitutions of Gly166 were produced.

Since hydrophobicity of aliphatic side-chains is directly proportional to side-chain surface area (Rose, G.D., et al. (1985) Science 229, 834-838; Reynolds, J.A., et al. (1974) Proc. Natl. Acad. Sci. USA 71, 2825-2927), increasing the hydrophobicity in the S-1 subsite may also sterically hinder binding of larger substrates. Because of difficulties in predicting the relative importance of these two opposing effects, we elected to generate twelve non-charged mutations at position 166 to determine the resulting specificities against non-charged substrates of varied size and hydrophobicity.

B. Cassette Mutagenesis of the P1 Binding Cleft

The preparation of mutant subtilisins containing the substitution of the hydrophobic amino acids Ala, Val and Phe into residue 166 has been described in EPO Publication No. 0130756. The same method was used to produce the remaining hydrophobic mutants at residue 166. In applying this method, two unique and silent restriction sites were introduced in the subtilisin genes to closely flank the target codon 166. As can be seen in Figure 13, the wild type sequence (line 1) was altered by site-directed mutagenesis in M13 using the indicated 37mer mutagenesis primer, to introduce a 13 bp deletion (dashedline) and unique SacI and XmaI sites (underlined sequences) that closely flank codon 166. The subtilisin gene fragment was subcloned back into the E. coli - B. subtilis shuttle plasmid, pBS42, giving the plasmid pΔ166 (Figure 13,

line 2). pΔ166 was cut open with *SacI* and *XmaI*, and gapped linear molecules were purified (Figure 13, line 3). Pools of synthetic oligonucleotides containing the mutation of interest were annealed to give duplex DNA cassettes that were ligated into gapped pΔ166 (underlined and overlined sequences in Figure 13, line 4). This construction restored the coding sequence except over position 166(NNN; line 4). Mutant sequences
 5 were confirmed by dideoxy sequencing. Asterisks denote sequence changes from the wild type sequence. Plasmids containing each mutant *B. amyloliquefaciens* subtilisin gene were expressed at roughly equivalent levels in a protease deficient strain of *B. subtilis*, BG2036 as previously described. EPO Publication No. 0130756; Yang, M., et al. (1984) *J. Bacteriol.* 160, 15-21; Estell, D.A., et al (1985) *J. Biol. Chem.* 260, 6518-6521.

10 C. Narrowing Substrate Specificity by Steric Hindrance

To probe the change in substrate specificity caused by steric alterations in the S-1 subsite, position 166 mutants were kinetically analyzed versus P1 substrates of increasing size (i.e., Ala, Met, Phe and Tyr).
 15 Ratios of k_{cat}/K_m are presented in log form in Figure 15 to allow direct comparisons of transition-state binding energies between various enzyme-substrate pairs.

According to transition state theory, the free energy difference between the free enzyme plus substrate ($E + S$) and the transition state complex ($E \cdot S^*$) can be calculated from equation (1),

$$(1) \quad \Delta G_T^\ddagger = -RT \ln k_{cat}/K_m + RT \ln kT/h$$

25 in which k_{cat} is the turnover number, K_m is the Michaelis constant, R is the gas constant, T is the temperature, k is Boltzmann's constant, and h is Planck's constant. Specificity differences are expressed quantitatively as differences between transition state binding energies (i.e., $\Delta\Delta G_T^\ddagger$), and can be calculated from equation (2).

$$(2) \quad \Delta\Delta G_T^\ddagger = -RT \ln (k_{cat}/K_m)_A / (k_{cat}/K_m)_B$$

30 A and B represent either two different substrates assayed against the same enzyme, or two mutant enzymes assayed against the same substrate.

As can be seen from Figure 15A, as the size of the side-chain at position 166 increases the substrate preference shifts from large to small P-1 side-chains. Enlarging the side-chain at position 166 causes k_{cat}/K_m to decrease in proportion to the size of the P-1 substrate side-chain (e.g., from Gly166 (wild-type)
 40 through W166, the k_{cat}/K_m for the Tyr substrate is decreased most followed in order by the Phe, Met and Ala P-1 substrates).

Specific steric changes in the position 166 side-chain, such as the presence of a β -hydroxyl group, β - or γ -aliphatic branching, cause large decreases in k_{cat}/K_m for larger P1 substrates. Introducing a β -hydroxyl group in going from A166 (Figure 15A) to S166 (Figure 15B), causes an 8 fold and 4 fold reduction in
 45 k_{cat}/K_m for Phe and Tyr substrates, respectively, while the values for Ala and Met substrates are unchanged. Producing a β -branched structure, in going from S166 to T166, results in a drop of 14 and 4 fold in k_{cat}/K_m for Phe and Tyr, respectively. These differences are slightly magnified for V166 which is slightly larger and isosteric with T166. Enlarging the β -branched substituents from V166 to I166 causes a lowering of k_{cat}/K_m between two and six fold toward Met, Phe and Tyr substrates. Inserting a γ -branched
 50 structure, by replacing M166 (Figure 15A) with L166 (Figure 15B), produces a 5 fold and 18 fold decrease in k_{cat}/K_m for Phe and Tyr substrates, respectively. Aliphatic γ -branching appears to induce less steric hindrance toward the Phe P-1 substrate than β -branching, as evidenced by the 100 fold decrease in k_{cat}/K_m for the Phe substrate in going from L166 to I166.

Reductions in k_{cat}/K_m resulting from increases in side chain size in the S-1 subsite, or specific
 55 structural features such as β - and γ -branching, are quantitatively illustrated in Figure 16. The k_{cat}/K_m values for the position 166 mutants determined for the Ala, Met, Phe, and Tyr P-1 substrates (top panel through bottom panel, respectively), are plotted versus the position 166 side-chain volumes (Chothia, C. (1984) *Ann. Rev. Biochem.* 53, 537-572). Catalytic efficiency for the Ala substrate reaches a maximum for

1166, and for the Met substrate it reaches a maximum between V166 and L166. The Phe substrate shows a broad kcat/Km peak but is optimal with A166. Here, the β -branched position 166 substitutions form a line that is parallel to, but roughly 50 fold lower in kcat/Km than side-chains of similar size [i.e., C166 versus T166, L166 versus I166]. The Tyr substrate is most efficiently utilized by wild type enzyme (Gly166), and there is a steady decrease as one proceeds to large position 166 side-chains. The β -branched and γ -branched substitutions form a parallel line below the other non-charged substitutions of similar molecular volume.

The optimal substitution at position 166 decreases in volume with increasing volume of the P1 substrate [i.e., I166/Ala substrate, L166/Met substrate, A166/Phe substrate, Gly166/Tyr substrate]. The combined volumes for these optimal pairs may approximate the volume for productive binding in the S-1 subsite. For the optimal pairs, Gly166/Tyr substrate, A166/Phe substrate, L166/Met substrate, V166/Met substrate, and I166/Ala substrate, the combined volumes are 266,295,313,339 and 261 \AA^3 , respectively. Subtracting the volume of the peptide backbone from each pair (i.e., two times the volume of glycine), an average side-chain volume of $160 \pm 32 \text{\AA}^3$ for productive binding can be calculated.

The effect of volume, in excess to the productive binding volume, on the drop in transition-state binding energy can be estimated from the Tyr substrate curve (bottom panel, Figure 16), because these data, and modeling studies (Figure 2), suggest that any substitution beyond glycine causes steric repulsion. A best-fit line drawn to all the data ($r = 0.87$) gives a slope indicating a loss of roughly 3 kcal/mol in transition state binding energy per 100\AA^3 of excess volume. (100\AA^3 is approximately the size of a leucyl side-chain.)

D. Enhanced Catalytic Efficiency Correlates with Increasing Hydrophobicity of the Position 166 Substitution

Substantial increases in kcat/Km occur with enlargement of the position 166 side-chain, except for the Tyr P-1 substrate (Figure 16). For example, kcat/Km increases in progressing from Gly166 to I166 for the Ala substrate (net of ten-fold), from Gly166 to L166 for the Met substrate (net of ten-fold) and from Gly166 to A166 for the Phe substrate (net of two-fold). The increases in kcat/Km cannot be entirely explained by the attractive terms in the van der Waals potential energy function because of their strong distance dependence ($1/r^6$) and because of the weak nature of these attractive forces (Jencks, W.P., Catalysis in Chemistry and Enzymology (McGraw-Hill, 1969) pp. 321-436; Fersht, A., Enzyme Structure and Mechanism (Freeman, San Francisco, 1977) pp. 226-287; Levitt, M. (1976) J. Mol. Biol. **104**, 59-107). For example, Levitt (Levitt, M. (1976) J. Mol. Biol. **104**, 59-107) has calculated that the van der Waals attraction between two methionyl residues would produce a maximal interaction energy of roughly -0.2 kcal/mol. This energy would translate to only 1.4 fold increase in kcat/Km.

The increases of catalytic efficiency caused by side-chain substitutions at position 166 are better accounted for by increases in the hydrophobicity of the S-1 subsite. The increase kcat/Km observed for the Ala and Met substrates with increasing position 166 side-chain size would be expected, because hydrophobicity is roughly proportional to side-chain surface area (Rose, G.D., et al. (1985) Science **229**, 834-838; Reynolds, J.A., et al. (1974) Proc. Natl. Acad. Sci. USA **71**, 2825-2927).

Another example that can be interpreted as a hydrophobic effect is seen when comparing kcat/Km for isosteric substitutions that differ in hydrophobicity such as S166 and C166 (Figure 16). Cysteine is considerably more hydrophobic than serine (-1.0 versus +0.3 kcal/mol) (Nozaki, Y., et al. (1971) J. Biol. Chem. **246**, 2211-2217; Tanford, C. (1978) Science **200**, 1012). The difference in hydrophobicity correlates with the observation that C166 becomes more efficient relative to Ser166 as the hydrophobicity of the substrates increases (i.e., Ala < Met < Tyr < Phe). Steric hindrance cannot explain these differences because serine is considerably smaller than cysteine (99 versus 118\AA^3). Paul, I.C., Chemistry of the -SH Group (ed. S. Patai, Wiley Interscience, New York, 1974) pp. 111-149.

E. Production of an Elastase-Like Specificity in Subtilisin

The I166 mutation illustrates particularly well that large changes in specificity can be produced by altering the structure and hydrophobicity of the S-1 subsite by a single mutation (Figure 17). Progressing through the small hydrophobic substrates, a maximal specificity improvement over wild type occurs for the Val substrate (16 fold in kcat/Km). As the substrate side chain size increases, these enhancements shrink to near unity (i.e., Leu and His substrates). The I166 enzyme becomes poorer against larger aromatic substrates of increasing size (e.g., I166 is over 1,000 fold worse against the Tyr substrate than is Gly166). We interpret the increase in catalytic efficiency toward the small hydrophobic substrates for I166 compared to Gly166 to the greater hydrophobicity of isoleucine (i.e., -1.8 kcal/mol versus 0). Nozaki, Y., et al. (1971) J. Biol. Chem. **246**, 2211-2217; Tanford, C. (1978) Science **200**, 1012. The decrease in catalytic efficiency

toward the very large substrates for I166 versus Gly166 is attributed to steric repulsion.

The specificity differences between Gly166 and I166 are similar to the specificity differences between chymotrypsin and the evolutionary relative, elastase (Harper, J.W., et al (1984) *Biochemistry* 23, 2995-3002). In elastase, the bulky amino acids, Thr and Val, block access to the P-1 binding site for large hydrophobic substrates that are preferred by chymotrypsin. In addition, the catalytic efficiencies toward small hydrophobic substrates are greater for elastase than for chymotrypsin as we observe for I166 versus Gly166 in subtilisin.

EXAMPLE 4

Substitution of Ionic Amino Acids for Gly166

The construction of subtilisin mutants containing the substitution of the ionic amino acids Asp, Asn, Gln, Lys and Arg are disclosed in EPO Publication No. 0130756. The present example describes the construction of the mutant subtilisin containing Glu at position 166 (E166) and presents substrate specificity data on these mutants. Further data on position 166 and 156 single and double mutants is presented *infra*.

pΔ166, described in Example 3, was digested with SacI and XmaI. The double strand DNA cassette (underlined and overlined) of line 4 in Figure 13 contained the triplet GAA for the codon 166 to encode the replacement of Glu for Gly166. This mutant plasmid designated pQ166 was propagated in BG2036 as described. This mutant subtilisin, together with the other mutants containing ionic substituent amino acids at residue 166, were isolated as described and further analyzed for variations in substrate specificity.

Each of these mutants was analyzed with the tetrapeptide substrates, succinyl-L-AlaL-AlaProL-X-p-nitroanilide, where X was Phe, Ala and Glu.

The results of this analysis are shown in Table IX.

TABLE IX

Position 166	P-1 Substrate (kcat/Km x 10 ⁻⁴)		
	Phe	Ala	Glu
Gly (wild type)	36.0	1.4	0.002
Asp (D)	0.5	0.4	<0.001
Glu (E)	3.5	0.4	<0.001
Asn (N)	18.0	1.2	0.004
Gln (Q)	57.0	2.6	0.002
Lys (K)	52.0	2.8	1.2
Arg (R)	42.0	5.0	0.08

These results indicate that charged amino acid substitutions at Gly166 have improved catalytic efficiencies (kcat/Km) for oppositely charged P-1 substrates (as much as 500 fold) and poorer catalytic efficiency for like charged P-1 substrates.

EXAMPLE 5

Substitution of Glycine at Position 169

The substitution of Gly169 in *B. amyloliquefaciens* subtilisin with Ala and Ser is described in EPO Publication No. 0130756. The same method was used to make the remaining 17 mutants containing all other substituent amino acids for position 169.

The construction protocol is summarized in Figure 18. The overscored and underscored double stranded DNA cassettes used contained the following triplet encoding the substitution of the indicated amino acid at residue 169.

5

10

GCT	A	ATG	M
TGT	C	AAC	N
GAT	D	CCT	P
GAA	E	CAA	Q
TTC	F	AGA	R
GGC	G	AGC	S
CAC	H	ACA	T
ATC	I	GTT	V
AAA	K	TGG	W
CTT	L	TAC	Y

Each of the plasmids containing a substituted Gly169 was designated pX169, where X represents the substituent amino acid. The mutant subtilisins were similarly designated.

15

Two of the above mutant subtilisins, A169 and S169, were analyzed for substrate specificity against synthetic substrates containing Phe, Leu, Ala and Arg in the P-1 position. The following results are shown in Table X.

TABLE X

20

Effect of Serine and Alanine Mutations at Position 169 on P-1 Substrate Specificity				
Position 169	P-1 Substrate [kcat/Km x 10 ⁻⁴]			
	Phe	Leu	Ala	Arg
Gly (wild type)	40	10	1	0.4
A169	120	20	1	0.9
S169	50	10	1	0.6

25

30

These results indicate that substitutions of Ala and Ser at Gly169 have remarkably similar catalytic efficiencies against a range of P-1 substrates compared to their position 166 counterparts. This is probably because position 169 is at the bottom of the P-1 specificity subsite.

35

EXAMPLE 6

Substitution at Position 104

40

Tyr104 has been substituted with Ala, His, Leu, Met and Ser. The method used was a modification of the site directed mutagenesis method. According to the protocol of Figure 19, a primer (shaded in line 4) introduced a unique HindIII site and a frame shift mutation at codon 104. Restriction-purification for the unique HindIII site facilitated the isolation of the mutant sequence (line 4). Restriction-selection against this HindIII site using primers in line 5 was used to obtain position 104 mutants.

45

The following triplets were used in the primers of Figure 19, line 5 for the 104 codon which substituted the following amino acids.

50

55

GCT	A	TTC	F
ATG	M	CCT	P
CTT	L	ACA	T
AGC	S	TGG	W
CAC	H	TAC	Y
CAA	Q	GTT	V
GAA	E	AGA	R
GGC	G	AAC	N
ATC	I	GAT	D
AAA	K	TGT	C

The substrates in Table XI were used to analyze the substrate specificity of these mutants. The results obtained for H104 subtilisin are shown in Table XI.

TABLE XI

Substrate	kcat		Km		Kcat/Km	
	WT	H104	WT	H104	WT	H104
sAAPFPNA	50.0	22.0	1.4×10^{-4}	7.1×10^{-4}	3.6×10^5	3.1×10^4
sAAPApNA	3.2	2.0	2.3×10^{-4}	1.9×10^{-3}	1.4×10^4	1×10^3
sFAPFPNA	26.0	38.0	1.8×10^{-4}	4.1×10^{-4}	1.5×10^5	9.1×10^4
sFAPApNA	0.32	2.4	7.3×10^{-5}	1.5×10^{-4}	4.4×10^3	1.6×10^4

From these data it is clear that the substitution of His for Tyr at position 104 produces an enzyme which is more efficient (higher kcat/Km) when Phe is at the P-4 substrate position than when Ala is at the P-4 substrate position.

EXAMPLE 7

Substitution of Ala152

Ala152 has been substituted by Gly and Ser to determine the effect of such substitutions on substrate specificity.

The wild type DNA sequence was mutated by the V152/P153 primer (Figure 20, line 4) using the above restriction-purification approach for the new KpnI site. Other mutant primers (shaded sequences Figure 20; S152, line 5 and G152, line 6) mutated the new KpnI site away and such mutants were isolated using the restriction-selection procedure as described above for loss of the KpnI site.

The results of these substitutions for the above synthetic substrates containing the P-1 amino acids Phe, Leu and Ala are shown in Table XII.

TABLE XII

Position 152	P-1 Substrate (kcat/Km $\times 10^{-4}$)		
	Phe	Leu	Ala
Gly (G)	0.2	0.4	<0.04
Ala (wild type)	40.0	10.0	1.0
Ser (S)	1.0	0.5	0.2

These results indicate that, in contrast to positions 166 and 169, replacement of Ala152 with Ser or Gly causes a dramatic reduction in catalytic efficiencies across all substrates tested. This suggests Ala152, at the top of the S-1 subsite, may be the optimal amino acid because Ser and Gly are homologous Ala substitutes.

EXAMPLE 8

Substitution at Position 156

Mutants containing the substitution of Ser and Gln for Glu156 have been constructed according to the overall method depicted in Figure 21. This method was designed to facilitate the construction of multiple mutants at position 156 and 166 as will be described hereinafter. However, by regenerating the wild type Gly166, single mutations at Glu156 were obtained.

The plasmid p Δ 166 is already depicted in line 2 of Figure 13. The synthetic oligonucleotides at the top right of Figure 21 represent the same DNA cassettes depicted in line 4 of Figure 13. The plasmid p166 in Figure 21 thus represents the mutant plasmids of Examples 3 and 4. In this particular example, p166 contains the wild type Gly166.

Construction of position 156 single mutants were prepared by ligation of the three fragments (1-3) indicated at the bottom of Figure 21. Fragment 3, containing the carboxy-terminal portion of the subtilisin gene including the wild type position 166 codon, was isolated as a 610 bp SacI-BamHI fragment. Fragment 1 contained the vector sequences, as well as the amino-terminal sequences of the subtilisin gene through codon 151. To produce fragment 1, a unique KpnI site at codon 152 was introduced into the wild type subtilisin sequence from pS4.5. Site-directed mutagenesis in M13 employed a primer having the sequence 5'-TA-GTC-GTT-GCG-GTA-CCC-GGT-AAC-GAA-3' to produce the mutation. Enrichment for the mutant sequence was accomplished by restriction with KpnI, purification and self ligation. The mutant sequence containing the KpnI site was confirmed by direct plasmid sequencing to give pV152. pV152 (~1 µg) was digested with KpnI and treated with 2 units of DNA polymerase I large fragment (Klenow fragment from Boehringer-Mannheim) plus 50 µM deoxynucleotide triphosphates at 37°C for 30 min. This created a blunt end that terminated with codon 151. The DNA was extracted with 1:1 volumes phenol and CHCl₃ and DNA in the aqueous phase was precipitated by addition of 0.1 volumes 5M ammonium acetate and two volumes ethanol. After centrifugation and washing the DNA pellet with 70% ethanol, the DNA was lyophilized. DNA was digested with BamHI and the 4.6kb piece (fragment 1) was purified by acrylamide gel electrophoresis followed by electroelution. Fragment 2 was a duplex synthetic DNA cassette which when ligated with fragments 1 and 3 properly restored the coding sequence except at codon 156. The top strand was synthesized to contain a glutamine codon, and the complementary bottom strand coded for serine at 156. Ligation of heterophosphorylated cassettes leads to a large and favorable bias for the phosphorylated over the non-phosphorylated oligonucleotide sequence in the final segregated plasmid product. Therefore, to obtain Q156 the top strand was phosphorylated, and annealed to the non-phosphorylated bottom strand prior to ligation. Similarly, to obtain S156 the bottom strand was phosphorylated and annealed to the non-phosphorylated top strand. Mutant sequences were isolated after ligation and transformation, and were confirmed by restriction analysis and DNA sequencing as before. To express variant subtilisins, plasmids were transformed into a subtilisin-neutral protease deletion mutant of *B. subtilis*, BG2036, as previously described. Cultures were fermented in shake flasks for 24 h at 37°C in LB media containing 12.5 mg/mL chloramphenicol and subtilisin was purified from culture supernatants as described. Purity of subtilisin was greater than 95% as judged by SDS PAGE.

These mutant plasmids designated pS156 and pQ156 and mutant subtilisins designated S156 and Q156 were analyzed with the above synthetic substrates where P-1 comprised the amino acids Glu, Gln, Met and Lys. The results of this analyses are presented in Example 9.

EXAMPLE 9

Multiple Mutants With Altered Substrate Specificity - Substitution at Positions 156 and 166

Single substitutions of position 166 are described in Examples 3 and 4. Example 8 describes single substitutions at position 156 as well as the protocol of Figure 21 whereby various double mutants comprising the substitution of various amino acids at positions 156 and 166 can be made. This example describes the construction and substrate specificity of subtilisin containing substitutions at position 156 and 166 and summarizes some of the data for single and double mutants at positions 156 and 166 with various substrates.

K166 is a common replacement amino acid in the 156/166 mutants described herein. The replacement of Lys for Gly166 was achieved by using the synthetic DNA cassette at the top right of Figure 21 which contained the triplet AAA for NNN. This produced fragment 2 with Lys substituting for Gly166.

The 156 substituents were Gln and Ser. The Gln and Ser substitutions at Gly156 are contained within fragment 3 (bottom right Figure 21).

The multiple mutants were produced by combining fragments 1, 2 and 3 as described in Example 8. The mutants Q156/K166 and S156/K166 were selectively generated by differential phosphorylation as described. Alternatively, the double 156/166 mutants, c.f. Q156/K166 and S156/K166, were prepared by ligation of the 4.6kb SacI-BamHI fragment from the relevant p156 plasmid containing the 0.6kb SacI-BamHI fragment from the relevant p166 plasmid.

These mutants, the single mutant K166, and the S156 and Q156 mutants of Example 8 were analyzed for substitute specificity against synthetic polypeptides containing Phe or Glu as the P-1 substrate residue. The results are presented in Table XIII.

TABLE XIII

Enzymes Compared (b)	Substrate		kcat	Km	kcat/Km	kcat/Km (mutant)	
	P-1 Residue					kcat/Km (wt)	
Glu156/Gly166 (WT)	Phe		50.00	1.4×10^{-4}	3.6×10^5	(1)	
K166	Glu		0.54	3.4×10^{-2}	1.6×10^1	(1)	
	Phe		20.00	4.0×10^{-5}	5.2×10^5	1.4	
Q156/K166	Glu		0.70	5.6×10^{-5}	1.2×10^4	750	
	Phe		30.00	1.9×10^{-5}	1.6×10^6	4.4	
S156/K166	Glu		1.60	3.1×10^{-5}	5.0×10^4	3100	
	Phe		30.00	1.8×10^{-5}	1.6×10^6	4.4	
S156	Glu		0.60	3.9×10^{-5}	1.6×10^4	1000	
	Phe		34.00	4.7×10^{-5}	7.3×10^5	2.0	
E156	Glu		0.40	1.8×10^{-3}	1.1×10^2	6.9	
	Phe		48.00	4.5×10^{-5}	1.1×10^6	3.1	
	Glu		0.90	3.3×10^{-3}	2.7×10^2	17	

As can be seen in Table XIV, either of these single mutations improve enzyme performance upon substrates with glutamate at the P-1 enzyme binding site. When these single mutations were combined, the resulting multiple enzyme mutants are better than either parent. These single or multiple mutations also alter the relative pH activity profiles of the enzymes as shown in Figure 23.

To isolate the contribution of electrostatics to substrate specificity from other chemical binding forces, these various single and double mutants were analyzed for their ability to bind and cleave synthetic substrates containing Glu, Gln, Met and Lys as the P-1 substrate amino acid. This permitted comparisons between side-chains that were more sterically similar but differed in charge (e.g., Glu versus Gln, Lys versus Met). Similarly, mutant enzymes were assayed against homologous P-1 substrates that were most sterically similar but differed in charge (Table XIV).

TABLE XIV
Kinetics of Position 156/166 Subtilisins
Determined for Different P1 Substrates

Enzyme Position (a)	Net Charge (b)	P-1 Substrate log kcat/Km (log 1/Km) (c)			
		Glu	Gln	Met	Lys
156 166					
Glu Asp	-2	n.d.	3.02 (2.56)	3.93 (2.74)	4.23 (3.00)
Glu Glu	-2	n.d.	3.06 (2.91)	3.86 (3.28)	4.48 (3.69)
Glu Asn	-1	1.62 (2.22)	3.85 (3.14)	4.99 (3.85)	4.15 (2.88)
Glu Gln	-1	1.20 (2.12)	4.36 (3.64)	5.43 (4.36)	4.10 (3.15)
Gln Asp	-1	1.30 (1.79)	3.40 (3.08)	4.94 (3.87)	4.41 (3.22)
Ser Asp	-1	1.23 (2.13)	3.41 (3.09)	4.67 (3.68)	4.24 (3.07)
Glu Met	-1	1.20 (2.30)	3.89 (3.19)	5.64 (4.83)	4.70 (3.89)
Glu Ala	-1	n.d.	4.34 (3.55)	5.65 (4.46)	4.90 (3.24)
Glu Gly (wt)	-1	1.20 (1.47)	3.85 (3.35)	5.07 (3.97)	4.60 (3.13)
Gln Gly	0	2.42 (2.48)	4.53 (3.81)	5.77 (4.61)	3.76 (2.82)
Ser Gly	0	2.31 (2.73)	4.09 (3.68)	5.61 (4.55)	3.46 (2.74)
Gln Asn	0	2.04 (2.72)	4.51 (3.76)	5.79 (4.66)	3.75 (2.74)
Ser Asn	0	1.91 (2.78)	4.57 (3.82)	5.72 (4.64)	3.68 (2.80)
Glu Arg	0	2.91 (3.30)	4.26 (3.50)	5.32 (4.22)	3.19 (2.80)
Glu Lys	0	4.09 (4.25)	4.70 (3.88)	6.15 (4.45)	4.23 (2.93)
Gln Lys	+1	4.70 (4.50)	4.64 (3.68)	5.97 (4.68)	3.23 (2.75)
Ser Lys	+1	4.21 (4.40)	4.84 (3.94)	6.16 (4.90)	3.73 (2.84)

Maximum difference:

log kcat/Km (log 1/Km) (d) 3.5 (3.0) 1.8 (1.4) 2.3 (2.2) -1.3 (-1.0)

Footnotes to Table XIV:

(a) *B. subtilis*, BG 2036, expressing indicated variant subtilisin were fermented and enzymes purified as previously described (Estell, *et al.* (1985) *J. Biol. Chem.* 260, 6518-6521). Wild type subtilisin is indicated (wt) containing Glu156 and Gly166.

(b) Net charge in the P-1 binding site is defined as the sum of charges from positions 156 and 166 at pH 8.6.

(c) Values for $k_{cat}(s^{-1})$ and $K_m(M)$ were measured in 0.1M Tris pH 8.6 at 25°C as previously described³ against P-1 substrates having the form succinyl-L-AlaL-AlaL-ProL-[X]-p-nitroanilide, where X is the indicated P-1 amino acid. Values for $\log 1/K_m$ are shown inside parentheses. All errors in determination of k_{cat}/K_m and $1/K_m$ are below 5%.

(d) Because values for Glu156/Asp166(D166) are too small to determine accurately, the maximum difference taken for GluP-1 substrate is limited to a charge range of +1 to -1 charge change.

n.d. = not determined

The k_{cat}/K_m ratios shown are the second order rate constants for the conversion of substrate to product, and represent the catalytic efficiency of the enzyme. These ratios are presented in logarithmic form to scale the data, and because $\log k_{cat}/K_m$ is proportional to the lowering of transition-state activation energy (ΔG^\ddagger). Mutations at position 156 and 166 produce changes in catalytic efficiency toward Glu, Gln, Met and Lys P-1 substrates of 3100, 60, 200 and 20 fold, respectively. Making the P-1 binding-site more positively charged [e.g., compare Gln156/Lys166 (Q156/K166) versus Glu156/Met166 (Glu156/M166)] dramatically increased k_{cat}/K_m toward the Glu P-1 substrate (up to 3100 fold), and decreased the catalytic efficiency toward the Lys P-1 substrate (up to 10 fold). In addition, the results show that the catalytic efficiency of wild type enzyme can be greatly improved toward any of the four P-1 substrates by mutagenesis of the P-1 binding site.

The changes in k_{cat}/K_m are caused predominantly by changes in $1/K_m$. Because $1/K_m$ is approximately equal to $1/K_s$, the enzyme-substrate association constant, the mutations primarily cause a change in substrate binding. These mutations produce smaller effects on k_{cat} that run parallel to the effects on $1/K_m$. The changes in k_{cat} suggest either an alteration in binding in the P-1 binding site in going from the Michaelis-complex E•S to the transition-state complex (E-S[‡]) as previously proposed (Robertus, J.D., *et al.* (1972) *Biochemistry* 11, 2439-2449; Robertus, J.D., *et al.* (1972) *Biochemistry* 11, 4293-4303), or change in the position of the scissile peptide bond over the catalytic serine in the E•S complex.

Changes in substrate preference that arise from changes in the net charge in the P-1 binding site show trends that are best accounted for by electrostatic effects (Figure 28). As the P-1 binding cleft becomes more positively charged, the average catalytic efficiency increases much more for the Glu P-1 substrate than for its neutral and isosteric P-1 homolog, Gln (Figure 28A). Furthermore, at the positive extreme both substrates have nearly identical catalytic efficiencies.

In contrast, as the P-1 site becomes more positively charged the catalytic efficiency toward the Lys P-1 substrate decreases, and diverges sharply from its neutral and isosteric homolog, Met (Figure 28B). The similar and parallel upward trend seen with increasing positive charge for the Met and Glu P-1 substrates probably results from the fact that all the substrates are succinylated on their amino-terminal end, and thus carry a formal negative charge.

The trends observed in $\log k_{cat}/K_m$ are dominated by changes in the K_m term (Figures 28C and 28D). As the pocket becomes more positively charged, the $\log 1/K_m$ values converge for Glu and Gln P-1 substrates (Figure 28C), and diverge for Lys and Met P-1 substrates (Figure 28D). Although less

pronounced effects are seen in log k_{cat} , the effects of P-1 charge on log k_{cat} parallel those seen in log $1/K_m$ and become larger as the P-1 pocket becomes more positively charged. This may result from the fact that the transition-state is a tetrahedral anion, and a net positive charge in the enzyme may serve to provide some added stabilization to the transition-state.

5 The effect of the change in P-1 binding-site charge on substrate preference can be estimated from the differences in slopes between the charged and neutral isosteric P-1 substrates (Figure 28B). The average change in substrate preference ($\Delta \log k_{cat}/K_m$) between charged and neutral isosteric substrates increases roughly 10-fold as the complementary charge or the enzyme increases (Table XV). When comparing Glu
10 versus Lys, this difference is 100-fold and the change in substrate preference appears predominantly in the K_m term.

TABLE XV

15	Differential Effect on Binding Site Charge on log k_{cat}/K_m or (log $1/K_m$) for P-1 Substrates that Differ in Charge ^(a)			
	Change in P-1 Binding Site Charge ^(b)	$\Delta \log k_{cat}/K_m$ ($\Delta \log 1/K_m$)		
		GluGln	MetLys	GluLys
20	-2 to -1	n.d.	1.2 (1.2)	n.d.
	-1 to 0	0.7 (0.6)	1.3 (0.8)	2.1 (1.4)
	0 to +1	1.5 (1.3)	0.5 (0.3)	2.0 (1.5)
	Avg. change in log k_{cat}/K_m or (log $1/K_m$) per unit charge change	1.1 (1.0)	1.0 (0.8)	2.1 (1.5)

25 ^(a) The difference in the slopes of curves were taken between the P-1 substrates over the charge interval given for log (k_{cat}/K_m) (Figure 28A, B) and (log $1/K_m$) (Figure 28C, D). Values represent the differential effect a charge change has in distinguishing the substrates that are compared.

^(b) Charge in P-1 binding site is defined as the sum of charges from positions 156 and 166.

30 The free energy of electrostatic interactions in the structure and energetics of salt-bridge formation depends on the distance between the charges and the microscopic dielectric of the media. To dissect these structural and microenvironmental effects, the energies involved in specific salt-bridges were evaluated. In addition to the possible salt-bridges shown (Figures 29A and 29B), reasonable salt-bridges can be built between a Lys P-1 substrate and Asp at position 166, and between a Glu P-1 substrate and a Lys at
35 position 166 (not shown). Although only one of these structures is confirmed by X-ray crystallography (Poulos, T.L., et al. (1976) *J. Mol. Biol.* 257 1097-1103), all models have favorable torsion angles (Sielecki, A.R., et al. (1979) *J. Mol. Biol.* 134, 781-804), and do not introduce unfavorable van der Waals contacts.

The change in charged P-1 substrate preference brought about by formation of the model salt-bridges above are shown in Table XVI.

TABLE XVI

Effect of Salt Bridge Formation Between Enzyme
and Substrate on P1 Substrate Preference (a)

Enzymes Compared (b)		Enzyme Position Changed	P-1 Substrates Compared	Substrate Preference $\Delta \log$ (kcat/Km)		Change in Substrate Preference $\Delta \Delta \log$ (kcat/Km)
1	2			1	2	
Glu156/Asp166	Gln156/Asp166	156	LysMet	+0.30	-0.53	0.83
Glu156/Asn166	Gln156/Asn166	156	LysMet	-0.84	-2.04	1.20
Glu156/Gly166	Gln156/Gly166	156	LysMet	-0.47	-2.10	1.63
Glu156/Lys-166	Gln156/Lys166	156	LysMet	-1.92	-2.74	0.82
				Ave $\Delta \Delta \log$ (kcat/Km)		1.10 \pm 0.3
Glu156/Asp166	Glu156/Asn166	166	LysMet	+0.30	-0.84	1.14
Glu156/Glu166	Glu156/Glu166	166	LysMet	+0.62	-1.33	1.95
Gln156/Asp166	Gln156/Asn166	166	LysMet	-0.53	-2.04	1.51
Ser156/Asp166	Ser156/Asn166	166	LysMet	-0.43	-2.04	1.61
Glu156/Lys166	Glu156/Met166	166	GluGln	-0.63	-2.69	2/06
				Ave $\Delta \Delta \log$ (kcat/Km)		1.70 \pm 0.3

Footnotes to Table XVI:

5 (a) Molecular modeling shows it is possible to form a salt bridge between the indicated charged P-1 substrate and a complementary charge in the P-1 binding site of the enzyme at the indicated position changed.

10 (b) Enzymes compared have sterically similar amino acid substitutions that differ in charge at the indicated position.

15 (c) The P-1 substrates compared are structurally similar but differ in charge. The charged P-1 substrate is complementary to the charge change at the position indicated between enzymes 1 and 2.

20 (d) Data from Table XIV was used to compute the difference in $\log(k_{cat}/K_m)$ between the charged and the non-charged P-1 substrate (i.e., the substrate preference). The substrate preference is shown separately for enzyme 1 and 2.

25 (e) The difference in substrate preference between enzyme 1 (more highly charged) and enzyme 2 (more neutral) represents the rate change accompanying the electrostatic interaction.

30 The difference between catalytic efficiencies (i.e., $\Delta \log k_{cat}/K_m$) for the charged and neutral P-1 substrates (e.g., Lys minus Met or Glu minus Gln) give the substrate preference for each enzyme. The change in substrate preference ($\Delta \Delta \log k_{cat}/K_m$) between the charged and more neutral enzyme homologs (e.g., Glu156/Gly166 minus Gln156(Q156)/Gly166) reflects the change in catalytic efficiency that may be
35 attributed solely to electrostatic effects.

These results show that the average change in substrate preference is considerably greater when electrostatic substitutions are produced at position 166 (50-fold in k_{cat}/K_m) versus position 156 (12-fold in k_{cat}/K_m). From these $\Delta \Delta \log k_{cat}/K_m$ values, an average change in transition-state stabilization energy can be calculated of -1.5 and -2.4 kcal/mol for substitutions at positions 156 and 166, respectively. This should
40 represent the stabilization energy contributed from a favorable electrostatic interaction for the binding of free enzyme and substrate to form the transition-state complex.

EXAMPLE 1045 Substitutions at Position 217

Tyr217 has been substituted by all other 19 amino acids. Cassette mutagenesis as described in EPO publication No. 0130756 was used according to the protocol of Figure 22. The EcoRV restriction site was used for restriction-purification of p Δ 217.

50 Since this position is involved in substrate binding, mutations here effect kinetic parameters of the enzyme. An example is the substitution of Leu for Tyr at position 217. For the substrate sAAPFpNa, this mutant has a k_{cat} of 277 s^{-1} and a K_m of 4.7×10^{-4} with a k_{cat}/K_m ratio of 6×10^5 . This represents a 5.5-fold increase in k_{cat} with a 3-fold increase in K_m over the wild type enzyme.

In addition, replacement of Tyr217 by Lys, Arg, Phe or Leu results in mutant enzymes which are more
55 stable at pHs of about 9-11 than the WT enzyme. Conversely, replacement of Tyr217 by Asp, Glu, Gly or Pro results in enzymes which are less stable at pHs of about 9-11 than the WT enzyme.

EXAMPLE 11

Multiple Mutants Having Altered Thermal Stability

- 5 B. amyloliquefacien subtilisin does not contain any cysteine residues. Thus, any attempt to produce thermal stability by Cys cross-linkage required the substitution of more than one amino acid in subtilisin with Cys. The following subtilisin residues were multiply substituted with cysteine:

Thr22/Ser87

Ser24/Ser87

- 10 Mutagenesis of Ser24 to Cys was carried out with a 5' phosphorylated oligonucleotide primer having the sequence

15 5'-pC-TAC-ACT-GGA-TGC-AAT-GTT-AAA-G-3'.

- (Asterisks show the location of mismatches and the underlined sequence shows the position of the altered Sau3A site.) The B. amyloliquefaciens subtilisin gene on a 1.5 kb EcoRI-BAMHI fragment from pS4.5 was cloned into M13mp11 and single stranded DNA was isolated. This template (M13mp11SUBT) was double primed with the 5' phosphorylated M13 universal sequencing primer and the mutagenesis primer. Adelman, et al. (1983) DNA 2, 183-193. The heteroduplex was transfected into competent JM101 cells and plaques were probed for the mutant sequence (Zoller, M.J., et al. (1982) Nucleic Acid Res. 10, 6487-6500; Wallace, et al. (1981) Nucleic Acid Res. 9, 3647-3656) using a tetramethylammonium chloride hybridization protocol (Wood, et al. (1985) Proc. Natl. Acad. Sci. USA 82, 1585-1588). The Ser87 to Cys mutation was prepared in a similar fashion using a 5' phosphorylated primer having the sequence

30 5'-pGGC-GTT-GCG-CCA-TGC-GCA-TCA-CT-3'.

- (The asterisk indicates the position of the mismatch and the underlined sequence shows the position of a new MstI site.) The C24 and C87 mutations were obtained at a frequency of one and two percent, respectively. Mutant sequences were confirmed by dideoxy sequencing in M13.

- 35 Mutagenesis of Tyr21/Thr22 to A21/C22 was carried out with a 5' phosphorylated oligonucleotide primer having the sequence

40 5'-pAC-TCT-CAA-GGC-GCT-TGT-GGC-TCA-AAT-GTT-3'.

- (The asterisks show mismatches to the wild type sequence and the underlined sequence shows the position of an altered Sau3A site.) Manipulations for heteroduplex synthesis were identical to those described for C24. Because direct cloning of the heteroduplex DNA fragment can yield increased frequencies of mutagenesis, the EcoRI-BamHI subtilisin fragment was purified and ligated into pBS42. E. coli MM 294 cells were transformed with the ligation mixture and plasmid DNA was purified from isolated transformants. Plasmid DNA was screened for the loss of the Sau3A site at codon 23 that was eliminated by the mutagenesis primer. Two out of 16 plasmid preparations had lost the wild type Sau3A site. The mutant sequence was confirmed by dideoxy sequencing in M13.

- Double mutants, C22/C87 and C24/C87, were constructed by ligating fragments sharing a common Clal site that separated the single parent cystine codons. Specifically, the 500 bp EcoRI-Clal fragment containing the 5' portion of the subtilisin gene (including codons 22 and 24) was ligated with the 4.7 kb Clal-EcoRI fragment that contained the 3' portion of the subtilisin gene (including codon 87) plus pBS42 vector sequence. E. coli MM 294 was transformed with ligation mixtures and plasmid DNA was purified from individual transformants. Double-cysteine plasmid constructions were identified by restriction site markers originating from the parent cysteine mutants (i.e., C22 and C24, Sau3A minus; Cys87, MstI plus). Plasmids from E. coli were transformed into B. subtilis BG2036. The thermal stability of these mutants as compared to wild type subtilisin are presented in Figure 30 and Tables XVII and XVIII.

TABLE XVII

Effect of DTT on the Half-Time of Autolytic Inactivation of Wild-Type and Disulfide Mutants of Subtilisin*			
Enzyme	t _{1/2}		-DTT/ + DTT
	-DDT	+ DTT	
	min		
Wild-type	95	85	1.1
C22/C87	44	25	1.8
C24/C87	92	62	1.5

(*) Purified enzymes were either treated or not treated with 25mM DTT and dialyzed with or without 10mM DTT in 2mM CaCl₂, 50mM Tris (pH 7.5) for 14 hr. at 4 °C. Enzyme concentrations were adjusted to 80μl aliquots were quenched on ice and assayed for residual activity. Half-times for autolytic inactivation were determined from semi-log plots of log₁₀ (residual activity) versus time. These plots were linear for over 90% of the inactivation.

TABLE XVIII

Effect of Mutations in Subtilisin on the Half-Time of Autolytic Inactivation at 58 °C*	
Enzyme	$t_{1/2}$
	min
Wild-type	120
C22	22
C24	120
C87	104
C22/C87	43
C24/C87	115

(*) Half-times for autolytic inactivation were determined for wild-type and mutant subtilisins as described in the legend to Table III. Unpurified and non-reduced enzymes were used directly from *B. subtilis* culture supernatants.

The disulfides introduced into subtilisin did not improve the autolytic stability of the mutant enzymes when compared to the wild-type enzyme. However, the disulfide bonds did provide a margin of autolytic stability when compared to their corresponding reduced double-cysteine enzyme. Inspection of a highly refined x-ray structure of wild-type *B. amyloliquefaciens* subtilisin reveals a hydrogen bond between Thr22 and Ser87. Because cysteine is a poor hydrogen donor or acceptor (Paul, I.C. (1974) in *Chemistry of the -SH Group* (Patai, S., ed.) pp. 111-149, Wiley Interscience, New York) weakening of 22/87 hydrogen bond may explain why the C22 and C87 single-cysteine mutant proteins are less autolytically stable than either C24 or wild-type (Table XVIII). The fact that C22 is less autolytically stable than C87 may be the result of the Tyr21A mutation (Table XVIII). Indeed, construction and analysis of Tyr21/C22 shows the mutant protein has an autolytic stability closer to that of C87. In summary, the C22 and C87 of single-cysteine mutations destabilize the protein toward autolysis, and disulfide bond formation increases the stability to a level less than or equal to that of wild-type enzyme.

EXAMPLE 12

Multiple Mutants Containing Substitutions at Position 222 and Position 166 or 169

Double mutants 166/222 and 169/222 were prepared by ligating together (1) the 2.3kb *A*cclI fragment from pS4.5 which contains the 5' portion of the subtilisin gene and vector sequences, (2) the 200bp *A*vallI fragment which contains the relevant 166 or 169 mutations from the respective 166 or 169 plasmids, and (3) the 2.2kb *A*vallI fragment which contains the relevant 222 mutation 3' and of the subtilisin genes and vector

sequence from the respective p222 plasmid.

Although mutations at position 222 improve oxidation stability they also tend to increase the K_m . An example is shown in Table XIX. In this case the A222 mutation was combined with the K166 mutation to give an enzyme with k_{cat} and K_m intermediate between the two parent enzymes.

TABLE XIX

	k_{cat}	K_m
WT	50	1.4×10^{-4}
A222	42	9.9×10^{-4}
K166	21	3.7×10^{-5}
K166/A222	29	2.0×10^{-4}
substrate sAAPFpNa		

EXAMPLE 13

Multiple Mutants Containing Substitutions at Positions 50, 156, 166, 217 and Combinations Thereof

The double mutant S156/A169 was prepared by ligation of two fragments, each containing one of the relevant mutations. The plasmid pS156 was cut with XmaI and treated with S1 nuclease to create a blunt end at codon 167. After removal of the nuclease by phenol/chloroform extraction and ethanol precipitation, the DNA was digested with BamHI and the approximately 4kb fragment containing the vector plus the 5' portion of the subtilisin gene through codon 167 was purified.

The pA169 plasmid was digested with KpnI and treated with DNA polymerase Klenow fragment plus 50 μ M dNTPs to create a blunt end codon at codon 168. The Klenow was removed by phenol/chloroform extraction and ethanol precipitation. The DNA was digested with BamHI and the 590bp fragment including codon 168 through the carboxy terminus of the subtilisin gene was isolated. The two fragments were then ligated to give S156/A169.

Triple and quadruple mutants were prepared by ligating together (1) the 220bp PvuII/HaeIII fragment containing the relevant 156, 166 and/or 169 mutations from the respective p156, p166 and/or p169 double of single mutant plasmid, (2) the 550bp HaeIII/BamHI fragment containing the relevant 217 mutant from the respective p217 plasmid, and (3) the 3.9kb PvuII/BamHI fragment containing the F50 mutation and vector sequences.

The multiple mutant F50/S156/A169/L217, as well as B. amyloliquefaciens subtilisin, B. licheniformis subtilisin and the single mutant L217 were analyzed with the above synthetic polypeptides where the P-1 amino acid in the substrate was Lys, His, Ala, Gln, Tyr, Phe, Met and Leu. These results are shown in Figures 26 and 27.

These results show that the F50/S156/A169/L217 mutant has substrate specificity similar to that of the B. licheniformis enzyme and differs dramatically from the wild type enzyme. Although only data for the L217 mutant are shown, none of the single mutants (e.g., F50, S156 or A169) showed this effect. Although B. licheniformis differs in 88 residue positions from B. amyloliquefaciens, the combination of only these four mutations accounts for most of the differences in substrate specificity between the two enzymes.

EXAMPLE 14

Subtilisin Mutants Having Altered Alkaline Stability

A random mutagenesis technique was used to generate single and multiple mutations within the B. amyloliquefaciens subtilisin gene. Such mutants were screened for altered alkaline stability. Clones having increased (positive) alkaline stability and decreased (negative) alkaline stability were isolated and sequenced to identify the mutations within the subtilisin gene. Among the positive clones, the mutants V107 and R213 were identified. These single mutants were subsequently combined to produce the mutant V107/R213.

One of the negative clones (V50) from the random mutagenesis experiments resulted in a marked decrease in alkaline stability. Another mutant (P50) was analyzed for alkaline stability to determine the effect

of a different substitution at position 50. The F50 mutant was found to have a greater alkaline stability than wild type subtilisin and when combined with the double mutant V107/R213 resulted in a mutant having an alkaline stability which reflected the aggregate of the alkaline stabilities for each of the individual mutants.

The single mutant R204 and double mutant C204/R213 were identified by alkaline screening after random cassette mutagenesis over the region from position 197 to 228. The C204/R213 mutant was thereafter modified to produce mutants containing the individual mutations C204 and R213 to determine the contribution of each of the individual mutations. Cassette mutagenesis using pooled oligonucleotides to substitute all amino acids at position 204, was utilized to determine which substitution at position 204 would maximize the increase in alkaline stability. The mutation from Lys213 to Arg was maintained constant for each of these substitutions at position 204.

A. Construction of pB0180, an *E. coli*-*B. subtilis* Shuttle Plasmid

The 2.9 kb EcoRI-BamHI fragment from pBR327 (Covarrubias, L., et al. (1981) Gene 13, 25-35) was ligated to the 3.7kb EcoRI-BamHI fragment of pBD64 (Gryczan, T., et al. (1980) J. Bacteriol., 141, 246-253) to give the recombinant plasmid pB0153. The unique EcoRI recognition sequence in pBD64 was eliminated by digestion with EcoRI followed by treatment with Klenow and deoxynucleotide triphosphates (Maniatis, T., et al. (eds.) (1982) in Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.). Blunt end ligation and transformation yielded pB0154. The unique AvaI recognition sequence in pB0154 was eliminated in a similar manner to yield pB0171. pB0171 was digested with BamHI and PvuII and treated with Klenow and deoxynucleotide triphosphates to create blunt ends. The 6.4 kb fragment was purified, ligated and transformed into LE392 cells (Enquist, L.W., et al. (1977) J. Mol. Biol. 111, 97-120), to yield pB0172 which retains the unique BamHI site. To facilitate subcloning of subtilisin mutants, a unique and silent KpnI site starting at codon 166 was introduced into the subtilisin gene from pS4.5 (Wells, J.A., et al. (1983) Nucleic Acids Res., 11, 7911-7925) by site-directed mutagenesis. The KpnI plasmid was digested with EcoRI and treated with Klenow and deoxynucleotide triphosphates to create a blunt end. The Klenow was inactivated by heating for 20 min at 68 °C, and the DNA was digested with BamHI. The 1.5 kb blunt EcoRI-BamHI fragment containing the entire subtilisin was ligated with the 5.8 kb NruI-BamHI from pB0172 to yield pB0180. The ligation of the blunt NruI end to the blunt EcoRI end recreated an EcoRI site. Proceeding clockwise around pB0180 from the EcoRI site at the 5' end of the subtilisin gene is the unique BamHI site at the 3' end of the subtilisin gene, the chloramphenicol and neomycin resistance genes and UB110 gram positive replication origin derived from pBD64, the ampicillin resistance gene and gram negative replication origin derived from pBR327.

B. Construction of Random Mutagenesis Library

The 1.5 kb EcoRI-BamHI fragment containing the *B. amyloliquefaciens* subtilisin gene (Wells et al., 1983) from pB0180 was cloned into M13mp11 to give M13mp11 SUBT essentially as previously described (Wells, J.A., et al. (1986) J. Biol. Chem., 261,6564-6570). Deoxyuridine containing template DNA was prepared according to Kunkel (Kunkel, T.A. (1985) Proc. Natl. Acad. Sci. USA, 82 488-492). Uridine containing template DNA (Kunkel, 1985) was purified by CsCl density gradients (Maniatis, T. et al. (eds.) (1982) in Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.). A primer (AvaI⁻) having the sequence

5' GAAAAAAGACCCTAGCGTCGCTTA

ending at codon -11, was used to alter the unique AvaI recognition sequence within the subtilisin gene. (The asterisk denotes the mismatches from the wild-type sequence and underlined is the altered AvaI site.)

The 5' phosphorylated AvaI primer (~320 pmol) and ~40 pmol (~120 µg) of uridine containing M13mp11 SUBT template in 1.88 ml of 53 mM NaCl, 7.4 mM MgCl₂ and 7.4 mM Tris.HCl (pH 7.5) were annealed by heating to 90 °C for 2 min. and cooling 15 min at 24 °C (Fig. 31). Primer extension at 24 °C was initiated by addition of 100 µL containing 1 mM in all four deoxynucleotide triphosphates, and 20 µL Klenow fragment (5 units/µL). The extension reaction was stopped every 15 seconds over ten min by addition of 10 µL 0.25 M EDTA (pH 8) to 50 µL aliquots of the reaction mixture. Samples were pooled, phenol chloroform extracted and DNA was precipitated twice by addition of 2.5 vol 100% ethanol, and washed twice with 70% ethanol.

The pellet was dried, and redissolved in 0.4 ml 1 mM EDTA, 10 mM Tris (pH 8).

Misincorporation of α -thiodeoxynucleotides onto the 3' ends of the pool of randomly terminated template was carried out by incubating four 0.2 ml solutions each containing one-fourth of the randomly terminated template mixture (~20 μ g), 0.25 mM of a given α -thiodeoxynucleotide triphosphate, 100 units AMV polymerase, 50 mM KCL, 10 mM MgCl₂, 0.4 mM dithiothreitol, and 50 mM Tris (pH 8.3) (Champoux, J.J. (1984) *Genetics*, 2, 454-464). After incubation at 37°C for 90 minutes, misincorporation reactions were sealed by incubation for five minutes at 37°C with 50 mM all four deoxynucleotide triphosphates (pH 8), and 50 units AMV polymerase. Reactions were stopped by addition of 25 mM EDTA (final), and heated at 68°C for ten min to inactivate AMV polymerase. After ethanol precipitation and resuspension, synthesis of closed circular heteroduplexes was carried out for two days at 14°C under the same conditions used for the timed extension reactions above, except the reactions also contained 1000 units T4 DNA ligase, 0.5 mM ATP and 1 mM β -mercaptoethanol. Simultaneous restriction of each heteroduplex pool with KpnI, BamHI, and EcoRI confirmed that the extension reactions were nearly quantitative. Heteroduplex DNA in each reaction mixture was methylated by incubation with 80 μ M S-adenosylmethionine and 150 units dam methylase for 1 hour at 37°C. Methylation reactions were stopped by heating at 68°C for 15 min.

One-half of each of the four methylated heteroduplex reactions were transformed into 2.5 ml competent *E. coli* JM101 (Messing, J. (1979) *Recombinant DNA Tech. Bull.*, 2, 43-48). The number of independent transformants from each of the four transformations ranged from 0.4-2.0 $\times 10^5$. After growing out phage pools, RF DNA from each of the four transformations was isolated and purified by centrifugation through CsCl density gradients. Approximately 2 μ g of RF DNA from each of the four pools was digested with EcoRI, BamHI and Aval. The 1.5 kb EcoRI-BamHI fragment (i.e., Aval resistant) was purified on low gel temperature agarose and ligated into the 5.5 kb EcoRI-BamHI vector fragment of pB0180. The total number of independent transformants from each α -thiodeoxynucleotide misincorporation plasmid library ranged from 1.2-2.4 $\times 10^4$. The pool of plasmids from each of the four transformations was grown out in 200 ml LB media containing 12.5 μ g/ml cmp and plasmid DNA was purified by centrifugation through CsCl density gradients.

C. Expression and Screening of Subtilisin Point Mutants

Plasmid DNA from each of the four misincorporation pools was transformed (Anagnostopoulos, C., et al. (1967), *J. Bacteriol.*, 81, 741-746) into BG2036. For each transformation, 5 μ g of DNA produced approximately 2.5 $\times 10^5$ independent BG2036 transformants, and liquid culture aliquots from the four libraries were stored in 10% glycerol at 70°C. Thawed aliquots of frozen cultures were plated on LB/5 μ g/ml cmp/1.6% skim milk plates (Wells, J.A., et al. (1983) *Nucleic Acids Res.*, 11, 7911-7925), and fresh colonies were arrayed onto 96-well microtiter plates containing 150 l per well LB media plus 12.5 μ g/ml cmp. After 1 h at room temperature, a replica was stamped (using a matched 96 prong stamp) onto a 132 mm BA 85 nitrocellulose filter (Schleicher and Scheull) which was layered on a 140 mm diameter LB/cmp/skim milk plate. Cells were grown about 16 h at 30°C until halos of proteolysis were roughly 5-7 mm in diameter and filters were transferred directly to a freshly prepared agar plate at 37°C containing only 1.6% skim milk and 50 mM sodium phosphate pH 11.5. Filters were incubated on plates for 3-6 h at 37°C to produce halos of about 5 mm for wild-type subtilisin and were discarded. The plates were stained for 10 min at 24°C with Coomassie blue solution (0.25% Coomassie blue (R-250) 25% ethanol) and destained with 25% ethanol, 10% acetic acid for 20 min. Zones of proteolysis appeared as blue halos on a white background on the underside of the plate and were compared to the original growth plate that was similarly stained and destained as a control. Clones were considered positive that produced proportionately larger zones of proteolysis on the high pH plates relative to the original growth plate. Negative clones gave smaller halos under alkaline conditions. Positive and negative clones were restreaked to colony purify and screened again in triplicate to confirm alkaline pH results.

D. Identification and Analysis of Mutant Subtilisins

Plasmid DNA from 5 ml overnight cultures of more alkaline active *B. subtilis* clones was prepared according to Birnboim and Doly (Birnboim, H.C., et al. (1979) *Nucleic Acid Res.* 7, 1513) except that incubation with 2 mg/ml lysozyme proceeded for 5 min at 37°C to ensure cell lysis and an additional phenol/CHCl₃ extraction was employed to remove contaminants. The 1.5 kb EcoRI-BamHI fragment containing the subtilisin gene was ligated into M13mp11 and template DNA was prepared for DNA sequencing (Messing, J., et al. (1982) *Gene*, 19 269-276). Three DNA sequencing primers ending at codon 26, +95, and +155 were synthesized to match the subtilisin coding sequence. For preliminary sequence

identification a single track of DNA sequence, corresponding to the dNTPas misincorporation library from which the mutant came, was applied over the entire mature protein coding sequence (i.e., a single dideoxyguanosine sequence track was applied to identify a mutant from the dGTPas library). A complete four track of DNA sequence was performed 200 bp over the site of mutagenesis to confirm and identify the mutant sequence (Sanger, F., et al., (1980) *J. Mol. Biol.*, **143**, 161-178). Confirmed positive and negative bacilli clones were cultured in LB media containing 12.5µg/mL cmp and purified from culture supernatants as previously described (Estell, D.A., et al. (1985) *J. Biol. Chem.*, **260**, 6518-6521). Enzymes were greater than 98% pure as analyzed by SDS-polyacrylamide gel electrophoresis (Laemmli, U.K. (1970), *Nature*, **227**, 680-685), and protein concentrations were calculated from the absorbance at 280 nm,

10

$$\epsilon_{280}^{0.1\%} = 1.17$$

15 (Maturbara, H., et al. (1965), *J. Biol. Chem.*, **240**, 1125-1130).

Enzyme activity was measured with 200µg/mL succinyl-L-AlaL-AlaL-ProL-Phep-nitroanilide (Sigma) in 0.1M Tris pH 8.6 or 0.1 M CAPS pH 10.8 at 25°C. Specific activity (µ moles product/min-mg) was calculated from the change in absorbance at 410 nm from production of p-nitroaniline with time per mg of enzyme (E410 = 8,480 M-lcm-l; Del Mar, E.G., et al. (1979), *Anal. Biochem.*, **99**, 316-320). Alkaline autolytic stability studies were performed on purified enzymes (200µg/mL) in 0.1 M potassium phosphate (pH 12.0) at 37°C. At various times aliquots were assayed for residual enzyme activity (Wells, J.A., et al. (1986) *J. Biol. Chem.*, **261**, 6564-6570).

25 E. Results

1. Optimization and analysis of mutagenesis frequency

A set of primer-template molecules that were randomly 3'-terminated over the subtilisin gene (Fig. 31) was produced by variable extension from a fixed 5'-primer (The primer mutated a unique *Ava*I site at codon 11 in the subtilisin gene). This was achieved by stopping polymerase reactions with EDTA after various times of extension. The extent and distribution of duplex formation over the 1 kb subtilisin gene fragment was assessed by multiple restriction digestion (not shown). For example, production of new *Hinf*I fragments identified when polymerase extension had proceeded past Ile110, Leu233, and Asp259 in the subtilisin gene.

35 Misincorporation of each dNTPas at randomly terminated 3' ends by AMV reverse transcriptase (Zakour, R.A., et al. (1982), *Nature*, **295**, 708-710; Zakour, R.A., et al. (1984), *Nucleic Acids Res.*, **12**, 6615-6628) used conditions previously described (Champoux, J.J., (1984), *Genetics*, **2**, 454-464). The efficiency of each misincorporation reaction was estimated to be greater than 80% by the addition of each dNTPas to the *Ava*I restriction primer, and analysis by polyacrylamide gel electrophoresis. Misincorporations were sealed by polymerization with all four dNTP's and closed circular DNA was produced by reaction with DNA ligase.

Several manipulations were employed to maximize the yield of the mutant sequences in the heteroduplex. These included the use of a deoxyuridine containing template (Kunkel, T.A. (1985), *Proc. Natl. Acad. Sci. USA*, **82** 488-492; Pukkila, P.J. et al. (1983), *Genetics*, **104**, 571-582), *in vitro* methylation of the mutagenic strand (Kramer, W. et al. (1982) *Nucleic Acids Res.*, **10** 6475-6485), and the use of *Ava*I restriction-selection against the wild-type template strand which contained a unique *Ava*I site. The separate contribution of each of these enrichment procedures to the final mutagenesis frequency was not determined, except that prior to *Ava*I restriction-selection roughly one-third of the segregated clones in each of the four pools still retained a wild-type *Ava*I site within the subtilisin gene. After *Ava*I restriction-selection greater than 98% of the plasmids lacked the wild-type *Ava*I site.

50 The 1.5 kb *Eco*RI-*Bam*HI subtilisin gene fragment that was resistant to *Ava*I restriction digestion, from each of the four CsCl purified M13 RF pools was isolated on low melting agarose. The fragment was ligated *in situ* from the agarose with a similarly cut *E. coli*-B. *subtilis* shuttle vector, pB0180, and transformed directly into *E. coli* LE392. Such direct ligation and transformation of DNA isolated from agarose avoided losses and allowed large numbers of recombinants to be obtained (>100,000 per µg equivalent of input M13 pool).

The frequency of mutagenesis for each of the four dNTPas misincorporation reactions was estimated from the frequency that unique restriction sites were eliminated (Table XX). The unique restriction sites

chosen for this analysis, Clal, PvuII, and KpnI, were distributed over the subtilisin gene starting at codons 35, 104, and 166, respectively. As a control, the mutagenesis frequency was determined at the PstI site located in the β lactamase gene which was outside the window of mutagenesis. Because the absolute mutagenesis frequency was close to the percentage of undigested plasmid DNA, two rounds of restriction-selection were necessary to reduce the background of surviving uncut wild-type plasmid DNA below the mutant plasmid (Table XX). The background of surviving plasmid from wild-type DNA probably represents the sum total of spontaneous mutations, uncut wild-type plasmid, plus the efficiency with which linear DNA can transform E. coli. Subtracting the frequency for unmutagenized DNA (background) from the frequency for mutant DNA, and normalizing for the window of mutagenesis sampled by a given restriction analysis (4-6 bp) provides an estimate of the mutagenesis efficiency over the entire coding sequence (~1000 bp).

15

20

25

30

35

40

45

50

55

TABLE XX

5	α -thiol dNTP misincor- porated ^(b)	Restriction Site Selection	% resistant clones ^c			% resistant clones over Background ^d	% mutants per 1000bp ^e
			1st round	2nd round	Total		
10	None	<u>PstI</u>	0.32	0.7	0.002	0	-
	G	<u>PstI</u>	0.33	1.0	0.003	0.001	0.2
	T	<u>PstI</u>	0.32	<0.5	<0.002	0	0
	C	<u>PstI</u>	0.43	3.0	0.013	0.011	3
15	None	<u>ClaI</u>	0.28	5	0.014	0	-
	G	<u>ClaI</u>	2.26	85	1.92	1.91	380
	T	<u>ClaI</u>	0.48	31	0.15	0.14	35
	C	<u>ClaI</u>	0.55	15	0.08	0.066	17
20	None	<u>PvuII</u>	0.08	29	0.023	0	-
	G	<u>PvuII</u>	0.41	90	0.37	0.35	88
	T	<u>PvuII</u>	0.10	67	0.067	0.044	9
	C	<u>PvuII</u>	0.76	53	0.40	0.38	95
25	None	<u>KpnI</u>	0.41	3	0.012	0	-
	G	<u>KpnI</u>	0.98	35	0.34	0.33	83
	T	<u>KpnI</u>	0.36	15	0.054	0.042	8
	C	<u>KpnI</u>	1.47	26	0.38	0.37	93
30	None	<u>KpnI</u>	0.41	3	0.012	0	-
	G	<u>KpnI</u>	0.98	35	0.34	0.33	83
	T	<u>KpnI</u>	0.36	15	0.054	0.042	8
	C	<u>KpnI</u>	1.47	26	0.38	0.37	93
35	None	<u>KpnI</u>	0.41	3	0.012	0	-
	G	<u>KpnI</u>	0.98	35	0.34	0.33	83
	T	<u>KpnI</u>	0.36	15	0.054	0.042	8
	C	<u>KpnI</u>	1.47	26	0.38	0.37	93

(a) Mutagenesis frequency is estimated from the frequency for obtaining mutations that alter unique restriction sites within the mutagenized subtilisin gene (i.e., ClaI, PvuII, or KpnI) compared to mutation frequencies of the PstI site, that is outside the window of mutagenesis.

(b) Plasmid DNA was from wild-type (none) or mutagenized by dNTPs misincorporation as described.

(c) Percentage of resistant clones was calculated from the fraction of clones obtained after three fold or greater over-digestion of the plasmid with the indicated restriction enzyme compared to a

non-digested control. Restriction-resistant plasmid DNA from the first round was subjected to a second round of restriction-selection. The total represents the product of the fractions of resistant clones obtained from both rounds of selection and gives percentage of restriction-site mutant clones in the original starting pool. Frequencies were derived from counting at least 20 colonies and usually greater than 100.

(d) Percent resistant clones was calculated by subtracting the percentage of restriction-resistant clones obtained for wild-type DNA (i.e., none) from that obtained for mutant DNA.

(e) This extrapolates from the frequency of mutation over each restriction site to the entire subtilisin gene (~1 kb). This has been normalized to the number of possible bases (4-6 bp) within each restriction site that can be mutagenized by a given misincorporation event.

From this analysis, the average percentage of subtilisin genes containing mutations that result from dGTP α s, dCTP α s, or dTTP α s misincorporation was estimated to be 90, 70, and 20 percent, respectively. These high mutagenesis frequencies were generally quite variable depending upon the dNTP α s and misincorporation efficiencies at this site. Misincorporation efficiency has been reported to be both dependent on the kind of mismatch, and the context of primer (Champoux, J.J., (1984); Skinner, J.A., et al. (1986) *Nucleic Acids Res.*, 14, 6945-6964). Biased misincorporation efficiency of dGTP α s and dCTP α s over dTTP α s has been previously observed (Shortle, D., et al. (1985), *Genetics*, 110, 539-555). Unlike the dGTP α s, dCTP α s, and dTTP α s libraries the efficiency of mutagenesis for the dATP α s misincorporation library could not be accurately assessed because 90% of the restriction-resistant plasmids analyzed simply lacked the subtilisin gene insert. This problem probably arose from self-ligation of the vector when the dATP α s mutagenized subtilisin gene was subcloned from M13 into pB0180. Correcting for the vector background, we estimate the mutagenesis frequency around 20 percent in the dATP α s misincorporation library. In a separate experiment (not shown), the mutagenesis efficiencies for dGTP α s and dTTP α s misincorporation were estimated to be around 50 and 30 percent, respectively, based on the frequency of reversion of an inactivating mutation at codon 169.

The location and identity of each mutation was determined by a single track of DNA sequencing corresponding to the misincorporated α thiodeoxynucleotide over the entire gene followed by a complete four track of DNA sequencing focused over the site of mutation. Of 14 mutants identified, the distribution was similar to that reported by Shortle and Lin (1985) except we did not observe nucleotide insertion or deletion mutations. The proportion of AG mutations was highest in the G misincorporation library, and some unexpected point mutations appeared in the dTTP α s and dCTP α s libraries.

2. Screening and Identification of Alkaline Stability Mutants of Subtilisin

It is possible to screen colonies producing subtilisin by halos of casein digestion (Wells, J.A. et al. (1983) *Nucleic Acids Res.*, 11, 7911-7925). However, two problems were posed by screening colonies under high alkaline conditions (>pH 11). First, *B. subtilis* will not grow at high pH, and we have been unable to transform an alkylphilic strain of bacillus. This problem was overcome by adopting a replica plating strategy in which colonies were grown on filters at neutral pH to produce subtilisin and filters subsequently transferred to casein plates at pH 11.5 to assay subtilisin activity. However, at pH 11.5 the casein micells no longer formed a turbid background and thus prevented a clear observation of proteolysis halos. The problem was overcome by briefly staining the plate with Coomassie blue to amplify proteolysis zones and acidifying the plates to develop casein micell turbidity. By comparison of the halo size produced on the reference growth plate (pH 7) to the high pH plate (pH 11.5), it was possible to identify mutant subtilisins that had increased (positives) or decreased (negatives) stability under alkaline conditions.

Roughly 1000 colonies were screened from each of the four misincorporation libraries. The percentage of colonies showing a differential loss of activity at pH 11.5 versus pH 7 represented 1.4, 1.8, 1.4, and 0.6% of the total colonies screened from the thiol dGTP α s, dATP α s, dTTP α s, and dCTP α s libraries, respectively. Several of these negative clones were sequenced and all were found to contain a single base change as expected from the misincorporation library from which they came. Negative mutants included A36, E170 and V50. Two positive mutants were identified as V107 and R213. The ratio of negatives to positives was roughly 50:1.

3. Stability and Activity of Subtilisin Mutants at Alkaline pH

Subtilisin mutants were purified and their autolytic stabilities were measured by the time course of inactivation at pH 12.0 (Figs. 32 and 33). Positive mutants identified from the screen (i.e., V107 and R213) were more resistant to alkaline induced autolytic inactivation compared to wild-type; negative mutants (i.e., E170 and V50) were less resistant. We had advantageously produced another mutant at position 50 (F50) by site-directed mutagenesis. This mutant was more stable than wild-type enzyme to alkaline autolytic inactivation (Fig. 33) At the termination of the autolysis study, SDS-PAGE analysis confirmed that each subtilisin variant had autolyzed to an extent consistent with the remaining enzyme activity.

The stabilizing effects of V107, R213, and F50 are cumulative. See Table XXI. The double mutant, V107/R213 (made by subcloning the 920 bp EcoRI-KpnI fragment of pB0180V107 into the 6.6 kb EcoRI-KpnI fragment of pB0180R213), is more stable than either single mutant. The triple mutant, F50/V107/R213 (made by subcloning the 735 bp EcoRI-PvuII fragment of pF50 (Example 2) into the 6.8 kb EcoRI-PvuII fragment of pB0180/V107, is more stable than the double mutant V107/R213 or F50. The inactivation curves show a biphasic character that becomes more pronounced the more stable the mutant analyzed. This may result from some destabilizing chemical modification(s) (eg., deamidation) during the autolysis study and/or reduced stabilization caused by complete digestion of larger autolysis peptides. These alkaline autolysis studies have been repeated on separately purified enzyme batches with essentially the same results. Rates of autolysis should depend both on the conformational stability as well as the specific activity of the subtilisin variant (Wells, J.A., et al. (1986), *J. Biol. Chem.*, 261, 6564-6570). It was therefore possible that the decreases in autolytic inactivation rates may result from decreases in specific activity of the more stable mutant under alkaline conditions. In general the opposite appears to be the case. The more stable mutants, if anything, have a relatively higher specific activity than wild-type under alkaline conditions and the less stable mutants have a relatively lower specific activity. These subtle effects on specific activity for V107/R213 and F50/V107/R213 are cumulative at both pH 8.6 and 10.8. The changes in specific activity may reflect slight differences in substrate specificity, however, it is noteworthy that only positions 170 and 107 are within 6A of a bound model substrate (Robertus, J.D., et al. (1972), *Biochemistry* 11, 2438-2449).

TABLE XXI

Relationship between relative specific activity at pH 8.6 or 10.8 and alkaline autolytic stability			
Enzyme	Relative specific activity		Alkaline autolysis half-time (min) ^b
	pH 8.6	pH 10.8	
Wild-type	100 \pm 1	100 \pm 3	86
Q170	46 \pm 1	28 \pm 2	13
V107	126 \pm 3	99 \pm 5	102
R213	97 \pm 1	102 \pm 1	115
V107/R213	116 \pm 2	106 \pm 3	130
V50	66 \pm 4	61 \pm 1	58
F50	123 \pm 3	157 \pm 7	131
F50/V107/R213	126 \pm 2	152 \pm 3	168

^(a) Relative specific activity was the average from triplicate activity determinations divided by the wild-type value at the same pH. The average specific activity of wild-type enzyme at pH 8.6 and 10.8 was 70 μ moles/min-mg and 37 μ moles/min-mg, respectively.

^(b) Time to reach 50% activity was taken from Figs. 32 and 33.

F. Random Cassette Mutagenesis of Residues 197 through 228

Plasmid pΔ222 (Wells, et al. (1985) Gene 34, 315-323) was digested with PstI and BamHI and the 0.4 kb PstI/BamHI fragment (fragment 1, see Fig. 34) purified from a polyacrylamide gel by electroelution.

5 The 1.5 kb EcoRI/BamHI fragment from pS4.5 was cloned into M13mp9. Site directed mutagenesis was performed to create the A197 mutant and simultaneously insert a silent SstI site over codons 195-196. The mutant EcoRI/BamHI fragment was cloned back into pBS42. The pA197 plasmid was digested with BamHI and SstI and the 5.3 kb BamHI/SstI fragment (fragment 2) was purified from low melting agarose.

10 Complimentary oligonucleotides were synthesized to span the region from SstI (codons 195-196) to PstI (codons 228-230). These oligodeoxynucleotides were designed to (1) restore codon 197 to the wild type, (2) re-create a silent KpnI site present in pΔ222 at codons 219-220, (3) create a silent SmaI site over codons 210-211, and (4) eliminate the PstI site over codons 228-230 (see Fig. 35). Oligodeoxynucleotides were synthesized with 2% contaminating nucleotides at each cycle of synthesis, e.g., dATP reagent was spiked with 2% dCTP, 2% dGTP, and 2% dTTP. For 97-mers, this 2% poisoning should give the following
15 percentages of non-mutant, single mutants and double or higher mutants per strand with two or more misincorporations per complimentary strand: 14% non-mutant, 28% single mutant, and 57% with ≥2 mutations, according to the general formula

$$20 \quad f = \frac{\mu^n}{n!} e^{-\mu} .$$

25 where μ is the average number of mutations and n is a number class of mutations and f is the fraction of the total having that number of mutations. Complimentary oligodeoxynucleotide pools were phosphorylated and annealed (fragment 3) and then ligated at 2-fold molar excess over fragments 1 and 2 in a three-way ligation.

E. coli MM294 was transformed with the ligation reaction, the transformation pool-grown up over night and the pooled plasmid DNA was isolated. This pool represented 3.4×10^4 independent transformants. This
30 plasmid pool was digested with PstI and then used to retransform E. coli. A second plasmid pool was prepared and used to transform B. subtilis (BG2036). Approximately 40% of the BG2036 transformants actively expressed subtilisin as judged by halo-clearing on casein plates. Several of the non-expressing transformants were sequenced and found to have insertions or deletions in the synthetic cassettes.
35 Expressing BG2036 mutants were arrayed in microtiter dishes with 150μl of LB/12.5μg/mL chloramphenicol (cmp) per well, incubated at 37 °C for 3-4 hours and then stamped in duplicate onto nitrocellulose filters laid on LB 1.5% skim milk/5μg/mL cmp plates and incubated overnight at 33 °C (until halos were approximately 4-8 mm in diameter). Filters were then lifted to stacks of filter paper saturated with 1 x Tide commercial grade detergent, 50 mM Na₂CO₃, pH 11.5 and incubated at 65 °C for 90 min. Overnight growth plates were
40 Commassie stained and destained to establish basal levels of expression. After this treatment, filters were returned to pH7/skim milk/20μg/mL tetracycline plates and incubated at 37 °C for 4 hours to overnight.

Mutants identified by the high pH stability screen to be more alkaline stable were purified and analyzed for autolytic stability at high pH or high temperature. The double mutant C204/R213 was more stable than wild type at either high pH or high temperature (Table XXII).

45 This mutant was dissected into single mutant parents (C204 and R213) by cutting at the unique SmaI restriction site (Fig. 35) and either ligating wild type sequence 3' to the SmaI site to create the single C204 mutant or ligating wild type sequence 5' to the SmaI site to create the single R213 mutant. Of the two single parents, C204 was nearly as alkaline stable as the parent double mutant (C04/R213) and slightly more thermally stable. See Table XXII. The R213 mutant was only slightly more stable than wild type under
50 both conditions (not shown).

Another mutant identified from the screen of the 197 to 228 random cassette mutagenesis was R204. This mutant was more stable than wild type at both high pH and high temperature but less stable than C204.

TABLE XXII

Stability of subtilisin variants

Purified enzymes (200 μ g/mL) were incubated in 0.1M phosphate, pH 12 at 30°C for alkaline autolysis, or in 2mM CaCl₂, 50mM MOPS, pH 7.0 at 62°C for thermal autolysis. At various times samples were assayed for residual enzyme activity. Inactivations were roughly pseudo-first order, and t 1/2 gives the time it took to reach 50% of the starting activity in two separate experiments.

<u>Subtilisin variant</u>	t 1/2 (alkaline autolysis)		t 1/2 (thermal autolysis)	
	Exp.	Exp.	Exp.	Exp.
	#1	#2	#1	#2
wild type	30	25	20	23
F50/V107/R213	49	41	18	23
R204	35	32	24	27
C204	43	46	38	40
C204/R213	50	52	32	36
L204/R213	32	30	20	21

G. Random Mutagenesis at Codon 204

Based on the above results, codon 204 was targeted for random mutagenesis. Mutagenic DNA cassettes (for codon at 204) all contained a fixed R213 mutation which was found to slightly augment the stability of the C204 mutant.

Plasmid DNA encoding the subtilisin mutant C204/R213 was digested with SstI and EcoRI and a 1.0 kb EcoRI/SstI fragment was isolated by electro-elution from polyacrylamide gel (fragment 1, see Fig. 35).

C204/R213 was also digested with SmaI and EcoRI and the large 4.7 kb fragment, including vector sequences and the 3' portion of coding region, was isolated from low melting agarose (fragment 2, see Fig. 36).

Fragments 1 and 2 were combined in four separate three-way ligations with heterophosphorylated fragments 3 (see Figs. 36 and 37). This heterophosphorylation of synthetic duplexes should preferentially drive the phosphorylated strand into the plasmid ligation product. Four plasmid pools, corresponding to the four ligations, were restricted with SmaI in order to linearize any single cut C204/R213 present from fragment 2 isolation, thus reducing the background of C204/R213. E. coli was then re-transformed with

Small-restricted plasmid pools to yield a second set of plasmid pools which are essentially free of C204/R213 and any non-segregated heterduplex material.

These second enriched plasmid pools were then used to transform *B. subtilis* (BG2036) and the resulting four mutant pools were screened for clones expressing subtilisin resistant to high pH/temperature inactivation. Mutants found positive by such a screen were further characterized and identified by sequencing.

The mutant L204/R213 was found to be slightly more stable than the wild type subtilisin. See Table XXII.

Having described the preferred embodiments of the present invention, it will appear to those ordinarily skilled in the art that various modifications may be made to the disclosed embodiments, and that such modifications are intended to be within the scope of the present invention.

Claims

1. A subtilisin mutant derived by the substitution of at least one amino acid residue of a precursor subtilisin with a different amino acid, so that the subtilisin mutant has at least one property which is different from the same property of the precursor subtilisin, characterized by the substitution at one or more of Tyr21, Thr22, Ser24, Asp36, Ala45, Gly46, Ala48, Ser49, Met50, Asn77, Ser87, Lys94, Val95, Leu96, Ile107, Gly110, Met124, Lys170, Tyr171, Pro172, Asp197, Met199, Ser204, Lys213, His67, Leu135, Gly97, Ser101, Gly102, Glu103, Gly127, Gly128, Pro129, Tyr214, and Gly215 of *Bacillus amyloliquefaciens* subtilisin and equivalent amino acid residues in other precursor subtilisins.
2. A subtilisin mutant having an amino acid sequence derived from the amino acid sequence of a precursor subtilisin by the substitution of more than one amino acid residue of said amino acid sequence of said precursor subtilisin by a different amino acid, so that the subtilisin mutant has at least one property which is different from the same property of the precursor subtilisin, characterized by substitutions at more than one of Tyr21, Thr22, Ser24, Asp32, Ser33, Asp36, Ala45, Ala48, Ser49, Met50, Ser87, Lys94, Val95, Tyr104, Ile107, Gly110, Met124, Ala152, Asn155, Glu156, Gly166, Gly169, Lys170, Tyr171, Pro172, Phe189, Asp197, Met199, Ser204, Lys213, Tyr217, Ser221, Met222, His67, Leu135, Gly97, Ser101, Gly102, Glu103, Gly127, Gly128, Pro129, Tyr214, and Gly215 of *Bacillus amyloliquefaciens* subtilisin and equivalent amino acid residues in other precursor subtilisins, with the proviso that when substitution is made at any residue in the group Asp32, Ser33, Tyr104, Ala152, Asn155, Glu156, Gly166, Gly169, Phe189, Tyr217 and Met222 a substitution is also made at at least one specified position not of that group.
3. The mutant of claim 2 wherein said combinations are selected from Thr22/Ser87, Ser24/Ser87, Ala45/Ala48, Ser49/Lys94, Ser49/Val95, Met50/Val95, Met50/Gly110, Met50/Met124, Met50/Met222, Met124/Met222, Tyr21/Thr22, Met50/Met124/Met222, Tyr21/Thr22/Ser87, Met50/Glu156/Gly166/Tyr217, Met50/Glu156/Tyr217, Ile170/Lys213, Ser204/Lys213, Met50/Ile107/Lys213 and Ser24/Met50/Ile107/Glu156/Gly166/Gly169/Ser204/Lys213/Gly215/Tyr217.
4. A subtilisin mutant derived by the deletion of one or more amino acid residues in a precursor subtilisin equivalent to 161-164 in *B. amyloliquefaciens* subtilisin, said deletion being made alone or in combination with substitutions in the amino acid sequence of the precursor subtilisin, and producing at least one property which is different from the same property of the precursor subtilisin.
5. A subtilisin mutant having altered substrate specificity when compared to a precursor subtilisin, the mutant being derived by the substitution of a different amino acid at the residue equivalent to Leu + 126 of *B. amyloliquefaciens* subtilisin, alone or in combination with other substitutions or deletions in the amino acid sequence of the precursor subtilisin.
6. A subtilisin mutant having altered substrate specificity when compared to a precursor subtilisin, the mutant being derived by the substitution of a different amino acid at the residue equivalent to Asp + 99 in *B. amyloliquefaciens* subtilisin, alone or in combination with other substitutions or deletions in the amino acid sequence of the precursor subtilisin.
7. A DNA sequence encoding the mutant of any one of the preceding claims.

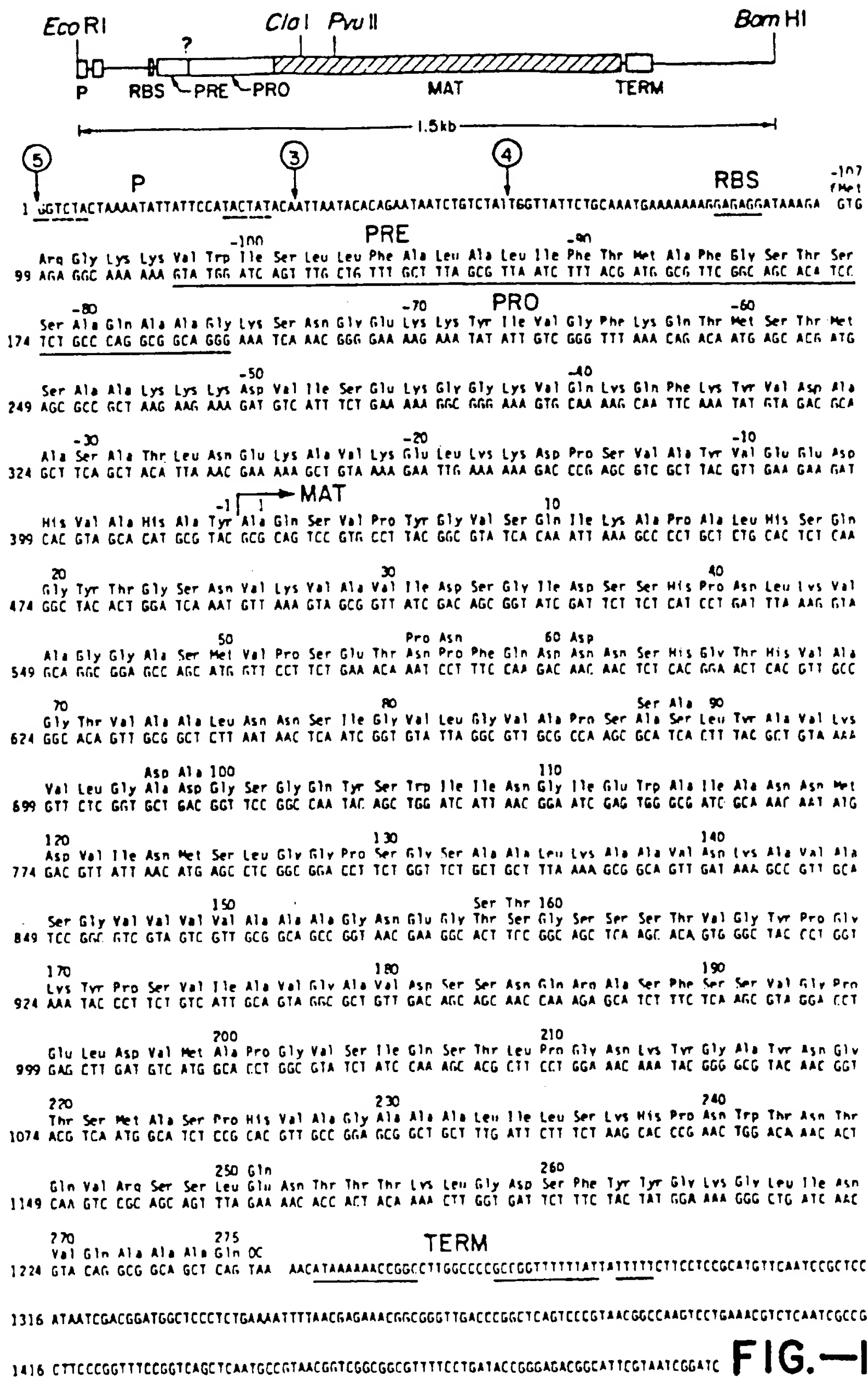
8. An expression vector containing the mutant DNA sequence of claim 7.
9. A host cell transformed with the expression vector of claim 8.

5 Patentansprüche

1. Subtilisinmutante, die durch Substitution zumindest eines Aminosäurerests eines Vorläufer-Subtilisins durch eine davon verschiedene Aminosäure hergeleitet ist, sodaß die Subtilisinmutante zumindest eine Eigenschaft aufweist, die sich von der gleichen Eigenschaft des Vorläufer-Subtilisins unterscheidet,
10 gekennzeichnet durch die Substitution an einem oder mehreren von Tyr21, Thr22, Ser24, Asp36, Ala45, Gly46, Ala48, Ser49, Met50, Asn77, Ser87, Lys94, Val95, Leu96, Ile107, Gly110, Met124, Lys170, Tyr171, Pro172, Asp197, Met199, Ser204, Lys213, His67, Leu135, Gly97, Ser101, Gly102, Glu103, Gly127, Gly128, Pro129, Tyr214 und Gly215 von Bacillus amyloliquefaciens-Subtilisin und äquivalenten Aminosäureresten in anderen Vorläufer-Subtilisinen.
- 15 2. Subtilisinmutante mit einer Aminosäuresequenz, die aus der Aminosäuresequenz eines Vorläufer-Subtilisins durch Substitution mehr als eines Aminosäurerests der Aminosäuresequenz des Vorläufer-Subtilisins durch eine davon verschiedene Aminosäure hergeleitet ist, sodaß die Subtilisinmutante zumindest eine Eigenschaft aufweist, die sich von der gleichen Eigenschaft des Vorläufer-Subtilisins unterscheidet, gekennzeichnet durch Substitutionen an mehr als einem von Tyr21, Thr22, Ser24,
20 Asp32, Ser33, Asp36, Ala45, Ala48, Ser49, Met50, Ser87, Lys94, Val95, Tyr104, Ile107, Gly110, Met124, Ala152, Asn155, Glu156, Gly166, Gly169, Lys170, Tyr171, Pro172, Phe189, Asp197, Met199, Ser204, Lys213, Tyr217, Ser221, Met222, His67, Leu135, Gly97, Ser101, Gly102, Glu103, Gly127, Gly128, Pro129, Tyr214 und Gly215 von Bacillus amyloliquefaciens-Subtilisin und äquivalenten Aminosäureresten in anderen Vorläufer-Subtilisinen, mit der Maßgabe, daß bei einer Substitution an irgendeinem Rest in der Gruppe Asp32, Ser33, Tyr104, Ala152, Asn155, Glu156, Gly166, Gly169, Phe189, Tyr217 und Met222 eine Substitution auch an zumindest einer bestimmten Position durchgeführt wird, die nicht dieser Gruppe angehört.
- 30 3. Mutante nach Anspruch 2, worin die Kombinationen aus Thr22/Ser87, Ser24/Ser87, Ala45/Ala48, Ser49/Lys94, Ser49/Val95, Met50/Val95, Met50/Gly110, Met50/Met124, Met50/Met222, Met124/Met222, Tyr21/Thr22, Met50/Met124/Met222, Tyr21/Tyr22/Ser87, Met50/Glu156/Gly166/Tyr217, Met50/Glu156/Tyr217, Ile170/Lys213, Ser204/Lys213, Met50/Ile107/Lys213 und Ser24/Met50/Ile107/Glu156/Gly166/Gly169/Ser204/Lys213/Gly215/Tyr217 ausgewählt sind.
- 35 4. Subtilisinmutante, die durch Löschung eines oder mehrerer Aminosäurereste in einem Vorläufer-Subtilisin, das 161-164 in B. amyloliquefaciens-Subtilisin äquivalent ist, hergeleitet ist, wobei die Löschung entweder alleine oder in Kombination mit Substitutionen in der Aminosäuresequenz des Vorläufer-Subtilisins erfolgt, und zumindest eine Eigenschaft ergibt, die sich von der gleichen Eigenschaft des Vorläufer-Subtilisins unterscheidet.
- 40 5. Subtilisinmutante mit geänderter Substratspezifität im Vergleich zu einem Vorläufer-Subtilisin, wobei die Mutante durch Substitution einer unterschiedlichen Aminosäure am Rest, der Leu + 126 von B. amyloliquefaciens-Subtilisin äquivalent ist, alleine oder in Kombination mit anderen Substitutionen oder
45 Löschungen in der Aminosäuresequenz des Vorläufer-Subtilisins hergeleitet ist.
6. Subtilisinmutante mit geänderter Substratspezifität im Vergleich zu einem Vorläufer-Subtilisin, wobei die Mutante durch Substitution einer unterschiedlichen Aminosäure am Rest, der Asp + 99 im B. amyloliquefaciens-Subtilisin äquivalent ist, alleine oder in Kombination mit anderen Substitutionen oder
50 Löschungen in der Aminosäuresequenz des Vorläufer-Subtilisins hergeleitet ist.
7. DNA-Sequenz, die für die Mutante nach einem der vorhergehenden Ansprüche kodiert.
8. Expressionsvektor, der die Mutanten-DNA-Sequenz von Anspruch 7 enthält.
- 55 9. Wirtszelle, die mit dem Expressionsvektor von Anspruch 8 transformiert ist.

Revendications

1. Mutant de subtilisine dérivé par la substitution d'au moins un résidu d'acide aminé d'une subtilisine précurseur et par un acide aminé différent de manière que le mutant de subtilisine ait au moins une propriété qui est différente de la même propriété de la subtilisine précurseur, caractérisé par la substitution à un ou plusieurs de Tyr21, Thr22, Ser24, Asp36, Ala45, Gly46, Ala48, Ser49, Met50, Asn77, Ser87, Lys94, Val95, Leu96, Ile107, Gly110, Met124, Lys170, Tyr171, Pro172, Asp197, Met199, Ser204, Lys213, His67, Leu135, Gly97, Ser101, Gly102, Glu103, Gly127, Gly128, Pro129, Tyr214 et Gly215 de la subtilisine de Bacillus amyloliquefaciens et les résidus d'acides aminés équivalents dans d'autres subtilisines précurseurs.
2. Mutant de subtilisine ayant une séquence d'acides aminés dérivée de la séquence d'acides aminés d'une subtilisine précurseur par la substitution de plus d'un résidu d'acide aminé de ladite séquence d'acides aminés de ladite subtilisine précurseur par un acide aminé différent de manière que le mutant de subtilisine ait au moins une propriété qui est différente de la même propriété de la subtilisine précurseur, caractérisé par des substitutions à plus d'un de Tyr21, Thr22, Ser24, Asp32, Ser33, Asp36, Ala45, Ala48, Ser49, Met50, Ser87, Lys94, Val95, Tyr104, Ile107, Gly110, Met124, Ala152, Asn155, Glu156, Gly166, Gly169, Lys170, Tyr171, Pro172, Phe189, Asp197, Met199, Ser204, Lys213, Tyr217, Ser221, Met222, His67, Leu135, Gly97, Ser101, Gly102, Glu103, Gly127, Gly128, Pro129, Tyr214 et Gly215 de la subtilisine de Bacillus amyloliquefaciens et des résidus d'acides aminés équivalents dans d'autres subtilisines précurseurs, à condition que quand la substitution est effectuée à tout résidu dans le groupe formé de Asp32, Ser33, Tyr104, Ala152, Asn155, Glu156, Gly166, Gly169, Phe189, Tyr217 et Met222, une substitution soit également effectuée en au moins une position spécifiée ne faisant pas partie de ce groupe.
3. Mutant de la revendication 2 où lesdites associations sont choisies parmi Thr22/Ser87, Ser24/Ser87, Ala45/Ala48, Ser49/Lys94, Ser49/Val95, Met50/Val95, Met50/Gly110, Met50/Met124, Met50/Met222, Met124/Met222, Tyr21/Thr22, Met50/Met124/Met222, Tyr21/Thr22/Ser87, Met50/Glu156/Gly166/Tyr217, Met50/Glu156/Tyr217, Ile170/Lys213, Ser204/Lys213, Met50/Ile107/Lys213 et Ser24/Met50/Ile107/Glu156/Gly166/Gly169/Ser204/Lys213/Gly215/Tyr217.
4. Mutant de subtilisine dérivé par la délétion d'un ou plusieurs résidus d'acides aminés dans une subtilisine précurseur équivalente à 161-164 dans la subtilisine de B. amyloliquefaciens, ladite délétion étant effectuée seule ou en association avec des substitutions dans la séquence d'acides aminés de la subtilisine précurseur et la production d'au moins une propriété qui est différente de la même propriété de la subtilisine précurseur.
5. Mutant de subtilisine ayant une spécificité modifiée du substrat en comparaison avec une subtilisine précurseur, le mutant étant dérivé par la substitution d'un acide aminé différent au résidu équivalent à Leu + 126 de la subtilisine de B. amyloliquefaciens, seule ou en association avec d'autres substitutions ou délétions dans la séquence d'acides aminés de la subtilisine précurseur.
6. Mutant de subtilisine ayant une spécificité modifiée de substrat en comparaison avec une subtilisine précurseur, le mutant étant dérivé par la substitution d'un acide aminé différent au résidu équivalent à Asp + 99 dans la subtilisine de B. amyloliquefaciens, seule ou en association avec d'autres substitutions ou délétions dans la séquence d'acides aminés de la subtilisine précurseur.
7. Séquence d'ADN codant le mutant selon l'une quelconque des revendications précédentes.
8. Vecteur d'expression contenant la séquence d'ADN du mutant de la revendication 7.
9. Cellule hôte transformée par le vecteur d'expression de la revendication 8.



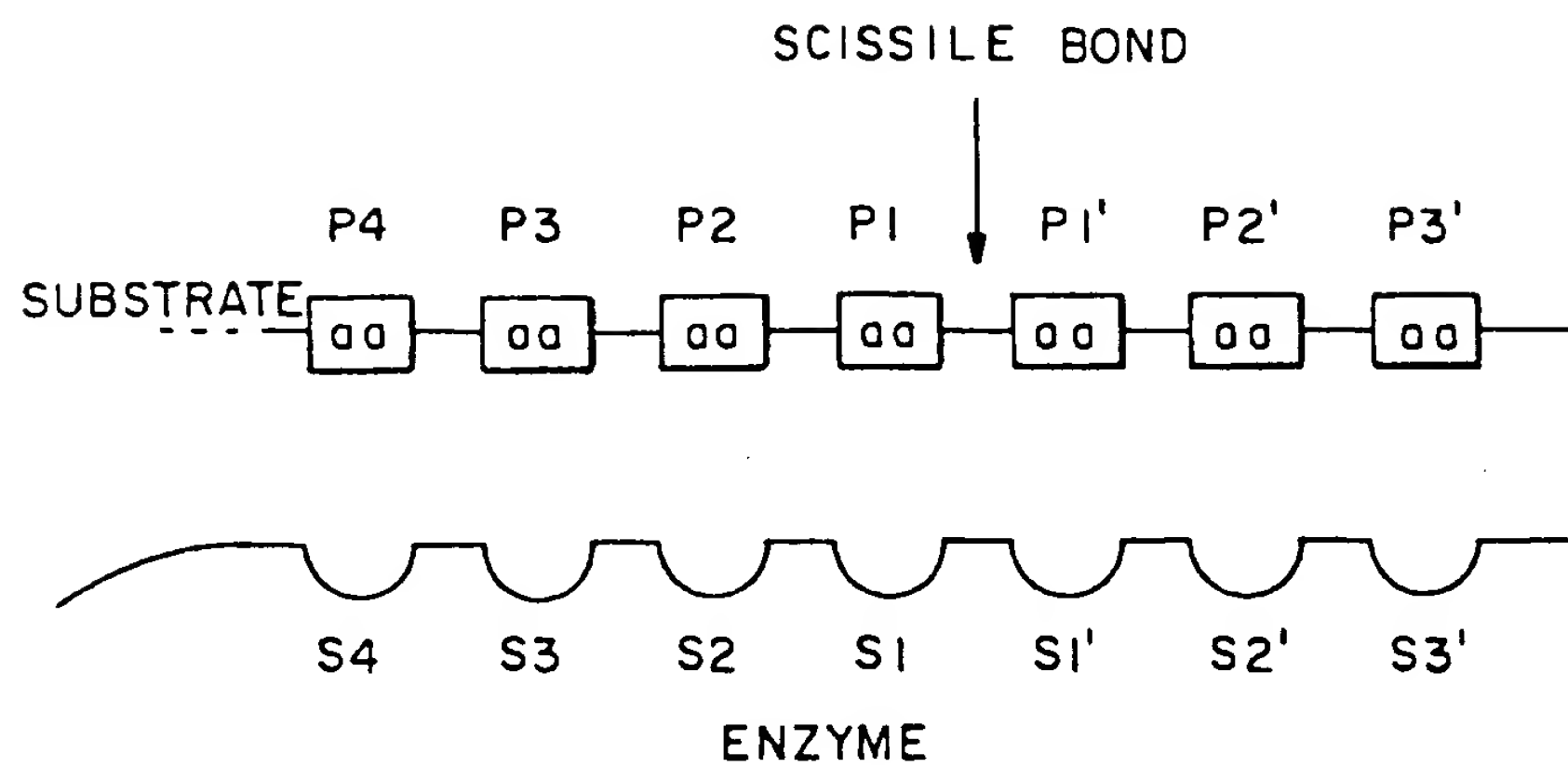


FIG.-2

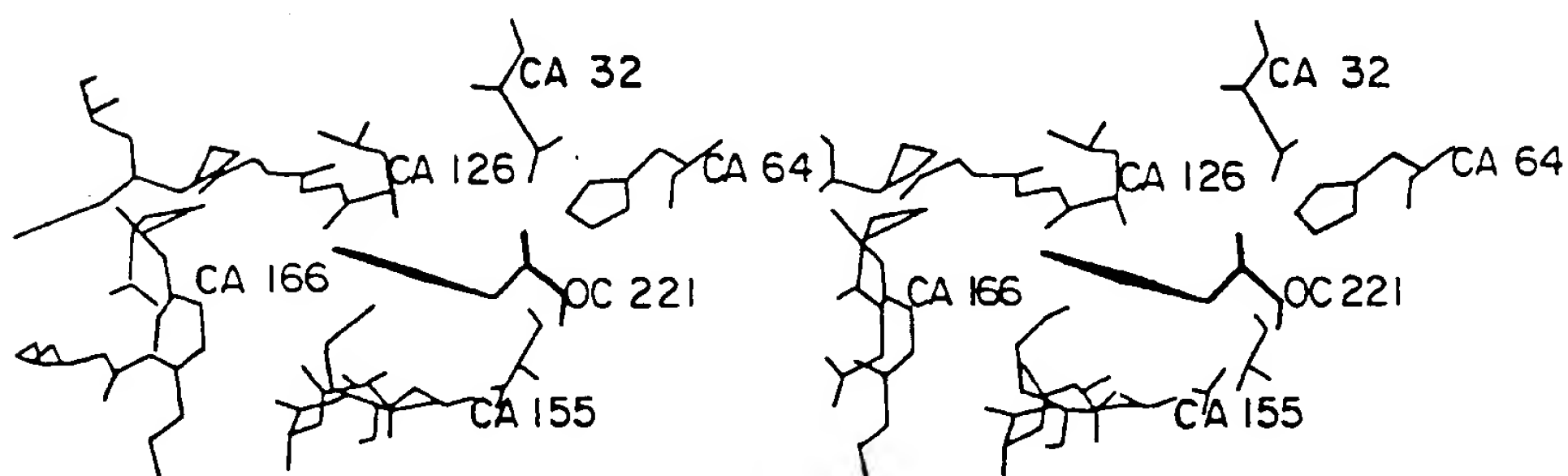


FIG.-3

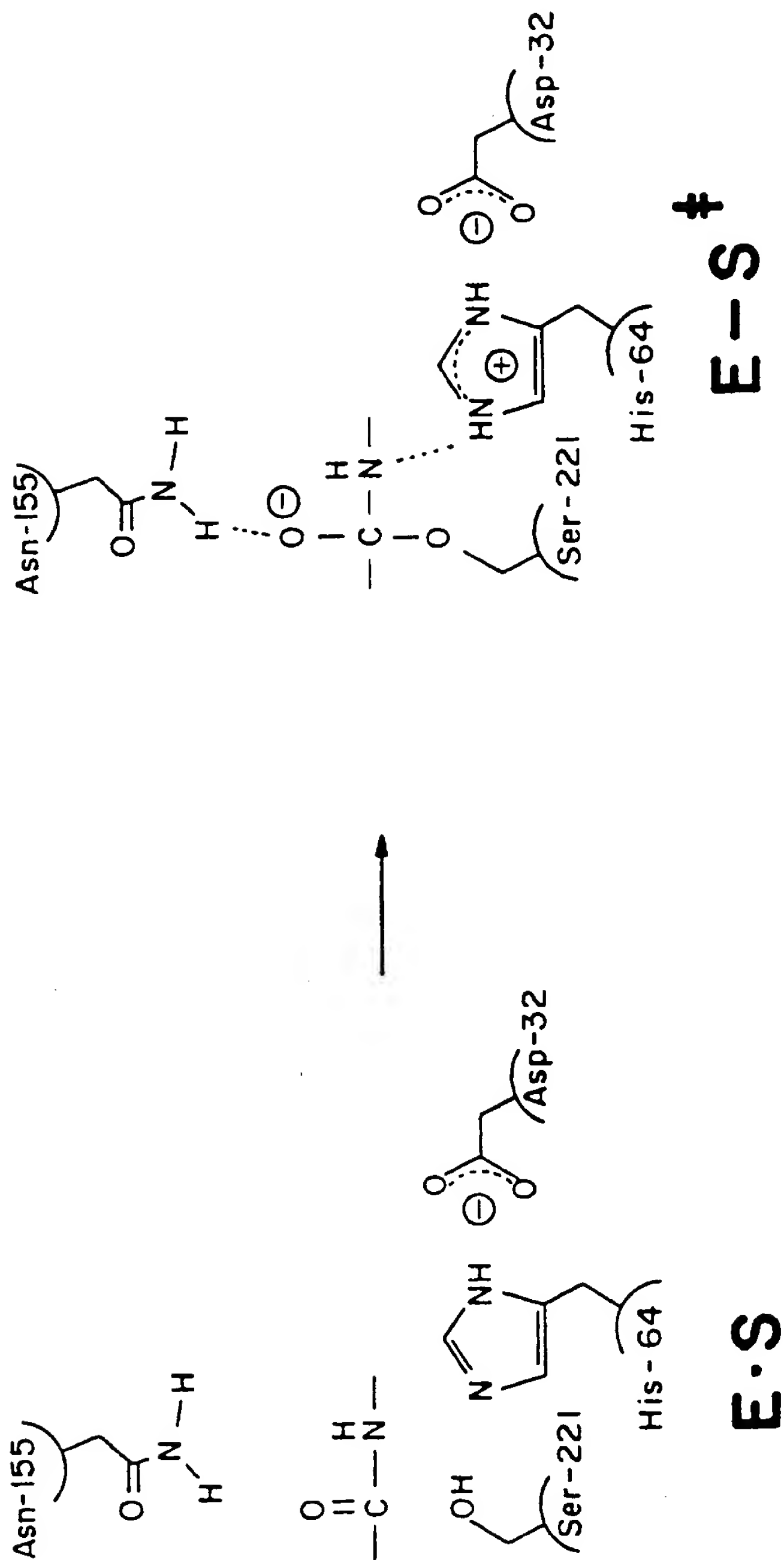


FIG.-4

Homology of *Bacillus* proteases

1. *Bacillus amyloliquifaciens*
2. *Bacillus subtilis* var. 1168
3. *Bacillus licheniformis* (carlsbergensis)

1									10									20
A	Q	S	V	P	Y	G	V	S	Q	I	K	A	P	A	L	H	S	Q
A	Q	S	V	P	Y	G	I	S	Q	I	K	A	P	A	L	H	S	Q
A	Q	T	V	P	Y	G	I	P	L	I	K	A	D	K	V	Q	A	Q
21									30									40
Y	T	G	S	N	V	K	V	A	V	I	D	S	G	I	D	S	S	H
Y	T	G	S	N	V	K	V	A	V	I	D	S	G	I	D	S	S	H
F	K	G	A	N	V	K	V	A	V	L	D	T	G	I	Q	A	S	H
41									50									60
D	L	K	V	A	G	G	A	S	H	V	P	S	E	T	N	P	F	Q
D	L	N	V	R	G	G	A	S	F	V	P	S	E	T	N	P	Y	Q
D	L	N	V	V	G	G	A	S	F	V	A	G	E	A	Y	N	T	.
61									70									80
N	N	S	H	G	T	H	V	A	G	T	V	A	A	L	N	N	S	I
G	S	S	H	G	T	H	V	A	G	T	I	A	A	L	N	N	S	I
G	N	G	H	G	T	H	V	A	G	T	V	A	A	L	D	N	T	T
81									90									100
V	L	G	V	A	P	S	A	S	L	Y	A	V	K	V	L	G	A	D
V	L	G	V	S	P	S	A	S	L	Y	A	V	K	V	L	D	S	T
V	L	G	V	A	P	S	V	S	L	Y	A	V	K	V	L	N	S	S
101									110									120
S	G	Q	Y	S	W	I	I	N	G	I	E	W	A	I	A	N	N	H
S	G	Q	Y	S	W	I	I	N	G	I	E	W	A	I	S	N	N	H
S	G	S	Y	S	G	I	V	S	G	I	E	W	A	T	T	N	G	H

FIG.—5A—1

121									130								140		
V	I	N	M	S	L	G	G	P	S	G	S	A	A	L	K	A	A	U	D
V	I	N	M	S	L	G	G	P	T	G	S	T	A	L	K	T	U	U	D
V	I	N	M	S	L	G	G	A	S	G	S	T	A	M	K	Q	A	U	D
141									150								160		
K	A	U	A	S	G	V	U	U	V	A	A	A	G	N	E	G	T	S	G
K	A	U	S	S	G	I	U	U	A	A	A	A	G	N	E	G	S	S	G
N	A	Y	A	R	G	U	U	U	U	A	A	A	G	N	S	G	N	S	G
161									170								180		
S	S	S	T	V	G	Y	P	G	K	Y	P	S	V	I	A	U	G	A	U
S	T	S	T	V	G	Y	P	A	K	Y	P	S	T	I	A	U	G	A	U
S	T	N	T	I	G	Y	P	A	K	Y	D	S	U	I	A	U	G	A	U
181									190								200		
D	S	S	N	Q	R	A	S	F	S	S	V	G	P	E	L	D	U	N	A
N	S	S	N	Q	R	A	S	F	S	S	A	G	S	E	L	D	U	N	A
D	S	N	S	N	R	A	S	F	S	S	U	G	A	E	L	E	U	N	A
201									210								220		
P	G	U	S	I	Q	S	T	L	P	G	N	K	Y	G	A	Y	N	G	T
P	G	U	S	I	Q	S	T	L	P	G	G	T	Y	G	A	Y	N	G	T
P	G	A	G	U	Y	S	T	Y	P	T	N	T	Y	A	T	L	N	G	T
221									230								240		
S	M	A	S	P	H	U	A	G	A	A	A	L	I	L	S	K	H	P	N
S	M	A	T	P	H	U	A	G	A	A	A	L	I	L	S	K	H	P	T
S	M	A	S	P	H	U	A	G	A	A	A	L	I	L	S	K	H	P	N
241									250								260		
W	T	N	T	Q	V	R	S	S	L	E	N	T	T	T	K	L	G	D	S
W	T	N	A	Q	V	R	D	R	L	E	S	T	A	T	Y	L	G	N	S
L	S	A	S	Q	U	R	N	R	L	S	S	T	A	T	Y	L	G	S	S
261									270										
F	Y	Y	G	K	G	L	I	N	U	Q	A	A	A	Q					
F	Y	Y	G	K	G	L	I	N	U	Q	A	A	A	Q					
F	Y	Y	G	K	G	L	I	N	U	E	A	A	A	Q					

FIG.—5A—2

ALIGNMENT OF D.AMYLOLIQUIFACIENS SUBTILISIN AND THERMITASE

1. B. mylioliquifactions subtilis in

2. thermite

[illegible]

FIG.—5B-1

A	A	A	G	N	E	S	T	S	150	S	S	S	T	U	B	Y	P	S	170
A	A	A	G	N	A	G	N	T	A	P	N	Y	P	A	K
Y	P	S	U	I	A	U	G	A	180	U	D	B	S	N	Q	R	A	S	190
Y	S	N	A	I	A	U	A	S	T	D	Q	N	D	N	K	S	S	F	S
S	U	G	P	E	L	D	U	M	200	A	P	G	U	S	I	Q	S	T	210
T	Y	G	S	U	U	D	U	A	A	P	B	S	W	I	Y	S	T	L	P
G	N	K	Y	G	A	J	N	G	220	T	S	M	A	S	P	H	U	A	230
T	S	T	Y	A	S	L	S	G	T	S	M	A	T	P	H	U	A	G	A
A	A	L	I	L	S	K	H	P	240	N	U	T	N	T	Q	U	R	S	250
A	G	L	L	A	S	Q	B	R	S	.	.	A	S	N	I	R	A	A	L
E	N	T	T	T	K	.	L	G	260	S	F	Y	Y	G	K	G	L	I	N
E	N	T	A	D	K	I	S	G	T	G	T	Y	W	A	K	B	R	U	N
270	U	Q	A	A	A	D													
A	Y	K	A	U	D	Y													

FIG.—5B—2

TOTALLY CONSERVED RESIDUES IN SUBTILISIN

[illegible]

FIG.—5C

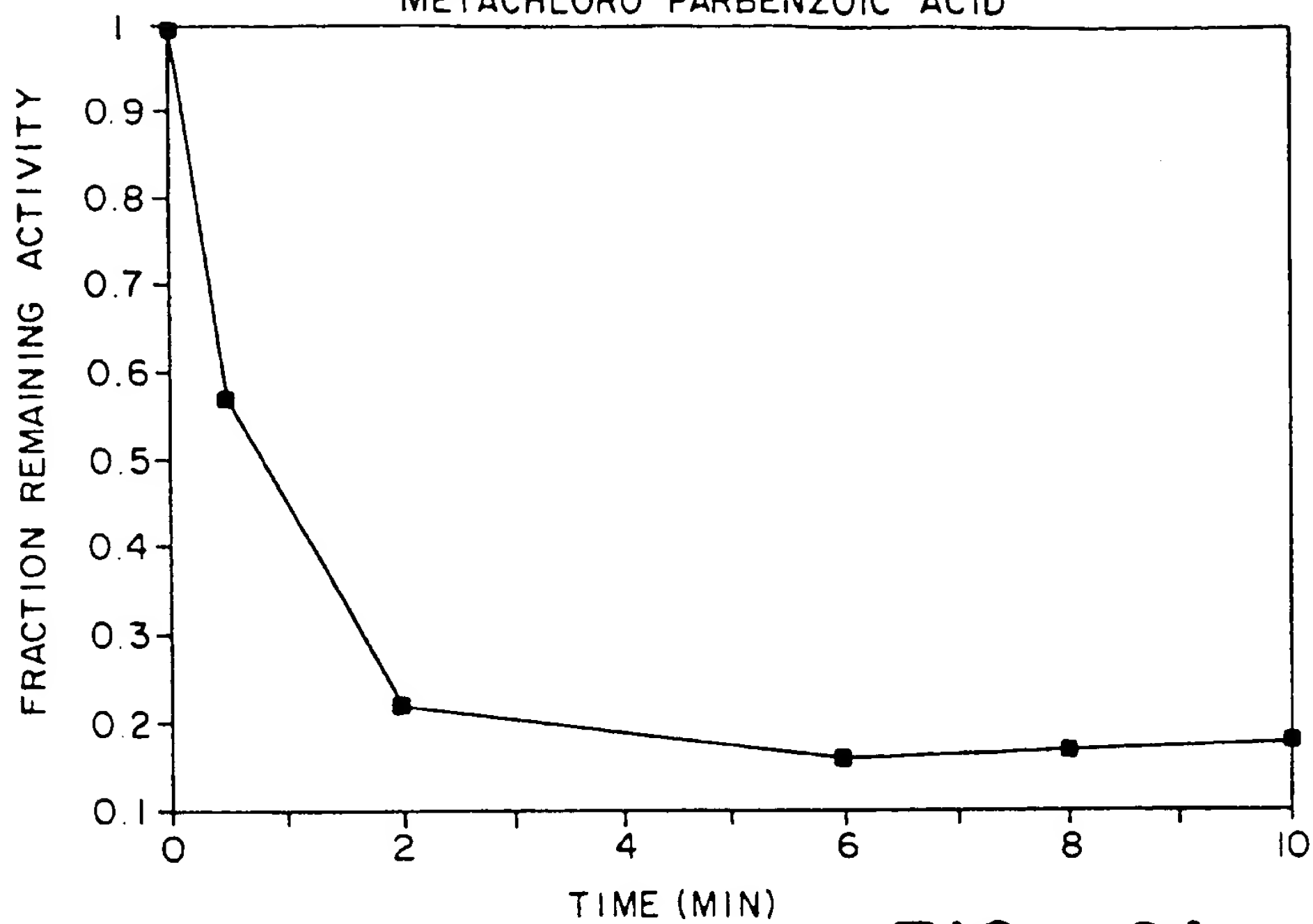
INACTIVATION OF L222 WITH
METACHLORO PARBENZOIC ACID

FIG.-6A

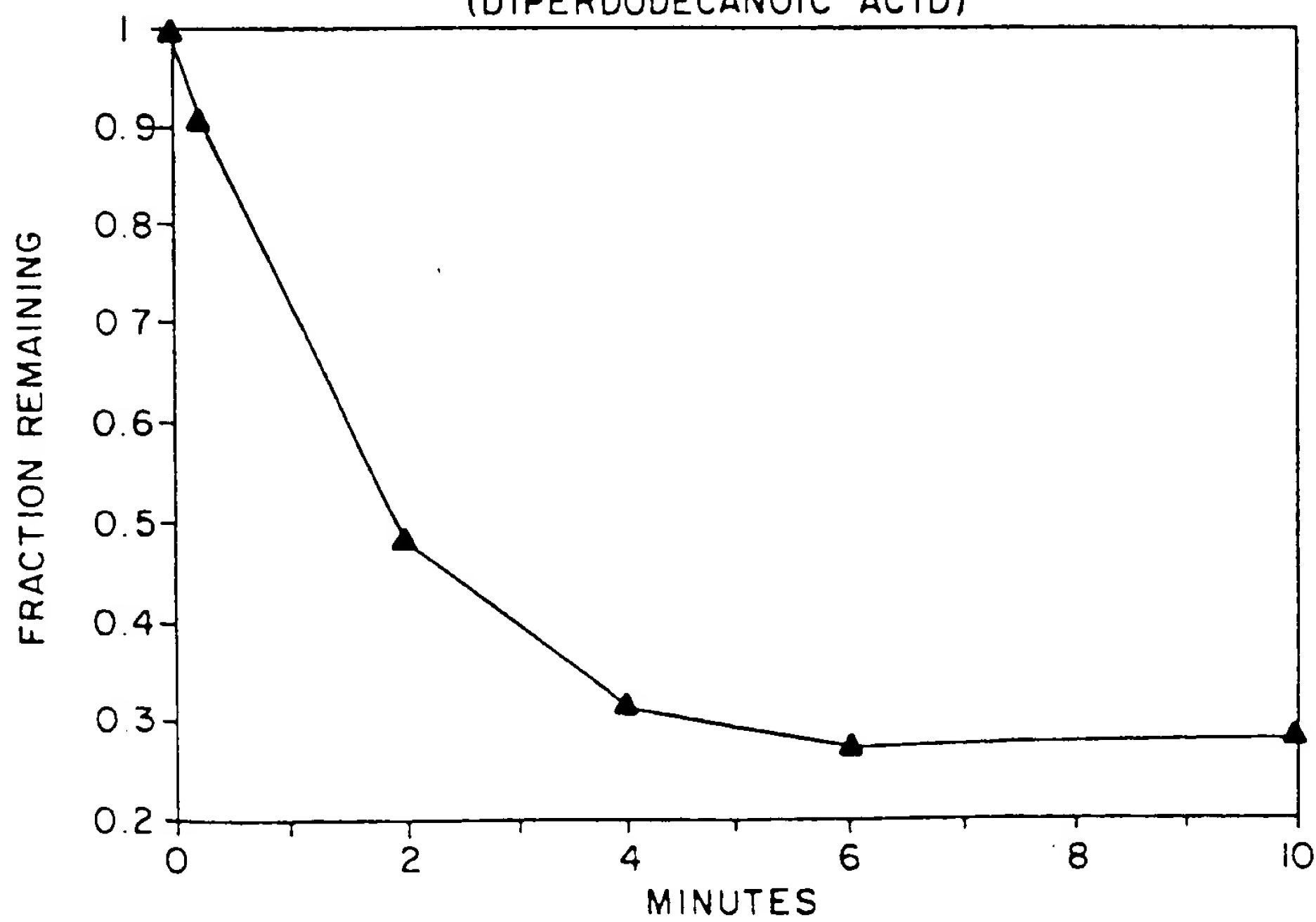
INACTIVATION OF Q222 BY DPDA
(DIPERDODECANOIC ACID)

FIG.-6B

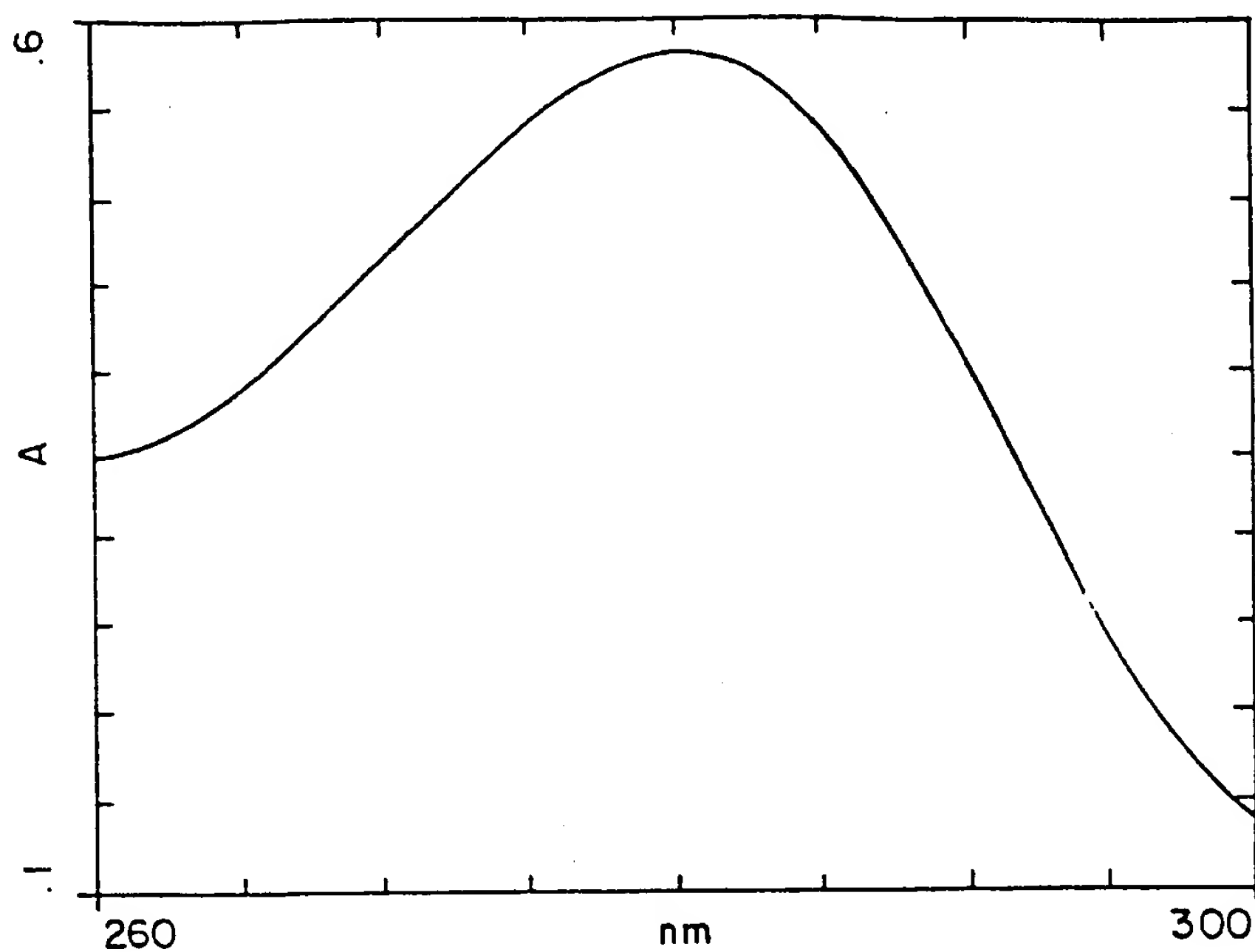


FIG. - 7A

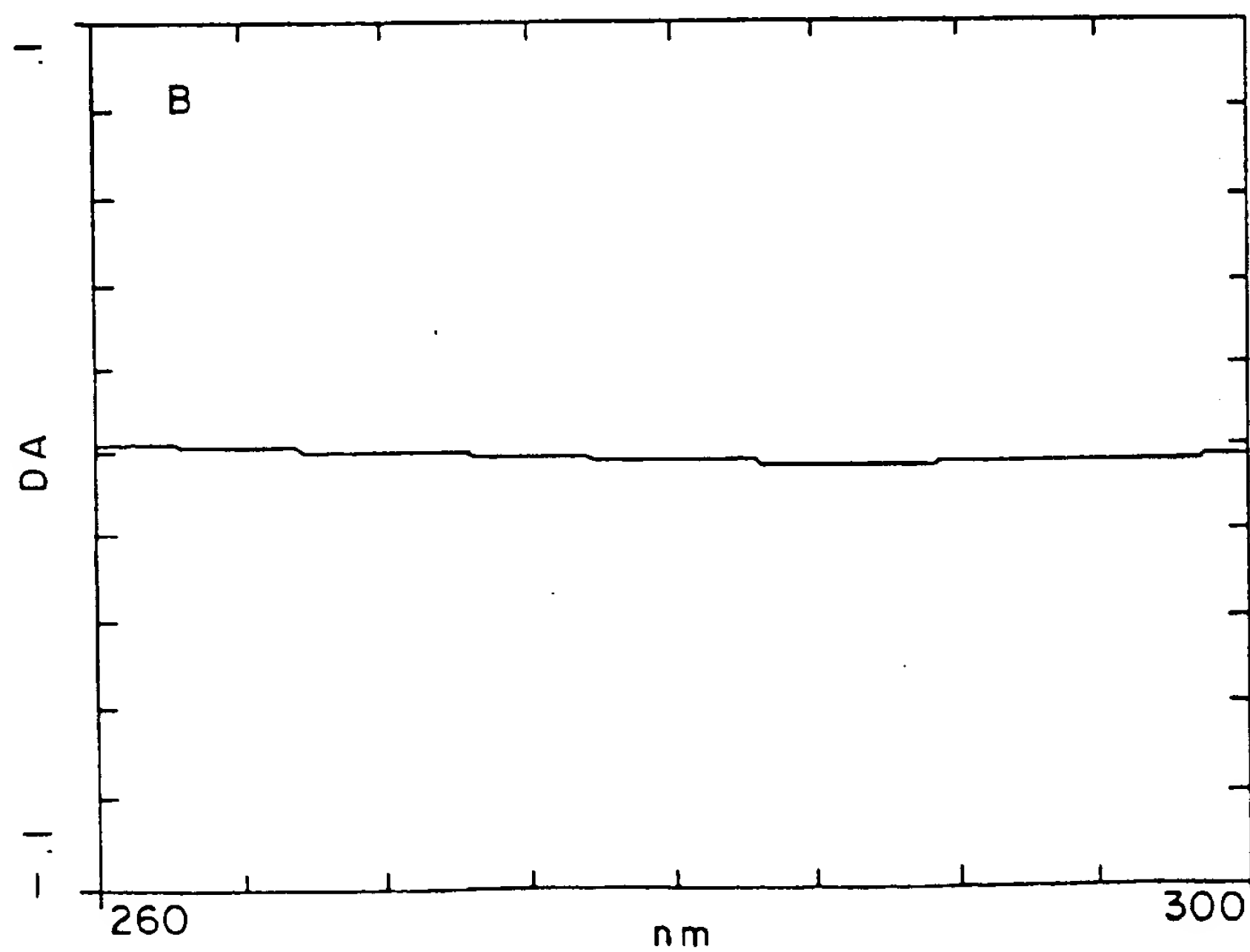


FIG. - 7B

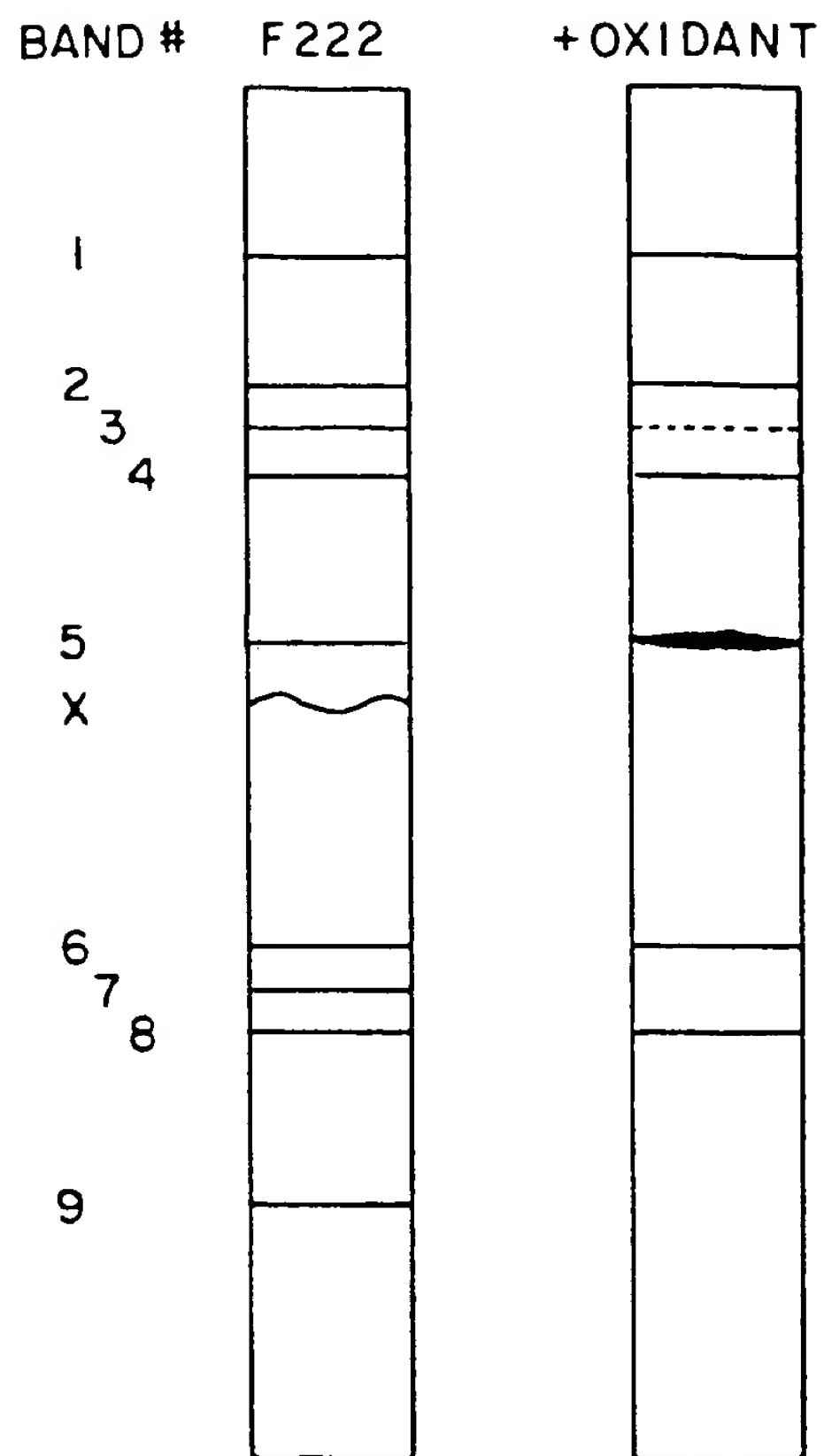


FIG.- 8

CNBr FRAGMENT MAP OF F222 MUTANT

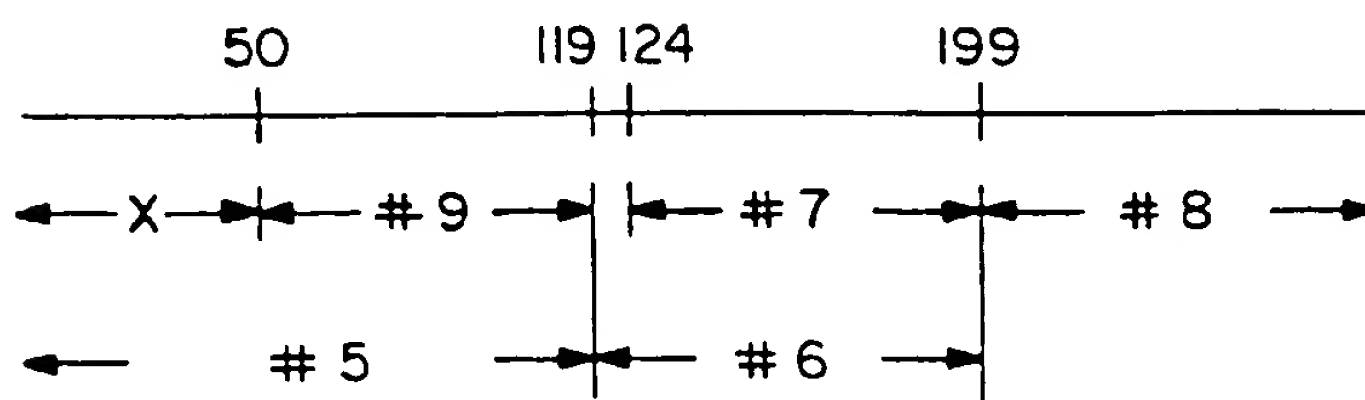


FIG.- 9

1. Codon number: 43 45
2. Wild type amino acid sequence: Lys-Val-Ala-Gly-Gly-Ala-Ser-Met-Val-Pro-Ser
3. Wild type DNA sequence: 5'-AAG-GTA-GCA-GGC-GGA-GCC-AGC-ATG-GTT-CCT-TCT
TTC-CAT-CGT-CCG-CCT-CCG-TCG-TAC-CAA-GGA-AGA-5'
4. pΔ50:

*** *

5'-AAG-GCC-T-----GC-ATG-GTA-CCT-TCT

TTC-CGG-A-----CG-TAC-CAT-GGA-AGA-5'

Su I *Kpn* I

*

PCT-TCT

CAT-GGA-AGA-5'
5. pΔ50 cut with *Stu* I/*Kpn* I

5'-AAG-G

TTC-Cp

*

PCT-TCT

CAT-GGA-AGA-5'
6. Cut pΔ50 ligated with cassettes:

5'-AAG-GTA-GCA-GGC-GGA-GCC-AGC-ATG-GTA-CCT-TCT

TCC-CAT-CGT-CCG-CCT-CCG-TCG-TAC-CAT-GGA-AGA-5'

*
7. Mutagenesis primer for pΔ50:

5'-CT-GAT-TTA-AAG-GCC-TGC-ATG-GTA-CCT-TCT-GA

*
8. Mutants made: V45, P45, V45/P48, E46, E48, V48, C49, C50, F50

FIG.—10

1. Codon number: 117 120 124 126 130
2. Wild type amino acid sequence: Asn-Asn-Met-Asp-Val-Ile-Asn-Met-Ser-Leu-Gly-Pro-Ser
3. Wild type DNA sequence: 5'-AAC-AAT-ATG-GAC-GTT-ATT-AAC-ATG-AGC-CTC-GGC-GGA-CCT-TCT
TTG-TTA-TAC-CTG-CAA-TAA-TTG-TAC-TCG-GAG-CCG-CCT-GGA-AGA-5'
4. pΔ124:

***	* * *	* * *
5'-AAC-AAT-ATG-GAT-ATC-----C-GGG-GGC-CCT-TCT		
TTG-TTA-TAC-CTA-TAG-----G-CCC-CCG-GGA-AGA-5'	Eco RV	Apa I
5. pΔ124 cut with Eco RV and Apa I

*	*	*
5'-AAC-AAT-ATG-GAT		PCT-TCT
TTG-TTA-TAC-CTAP		CCG-GGA-AGA-5'
6. Cut pΔ124 ligated with cassettes:

*	*	*
5'-AAC-AAT-ATG-GAT-GTT-ATT-AAC-ATG-AGC-CTC-GGC-GGC-CCT-TCT		
TTG-TTA-TAC-CTA-CAA-TAA-TTG-TAC-TCG-GAG-CCG-CCG-GGA-AGA-5'		
7. Mutagenesis primer for pΔ124::

* ** *	* * *	* * *
5'-AAC-AAT-ATG-GAT-ATC-C-GGG-GGC-CCT-TCT-GGT-TC-3'		
8. Mutants made: I 124, L 124 AND C126

FIG.—II

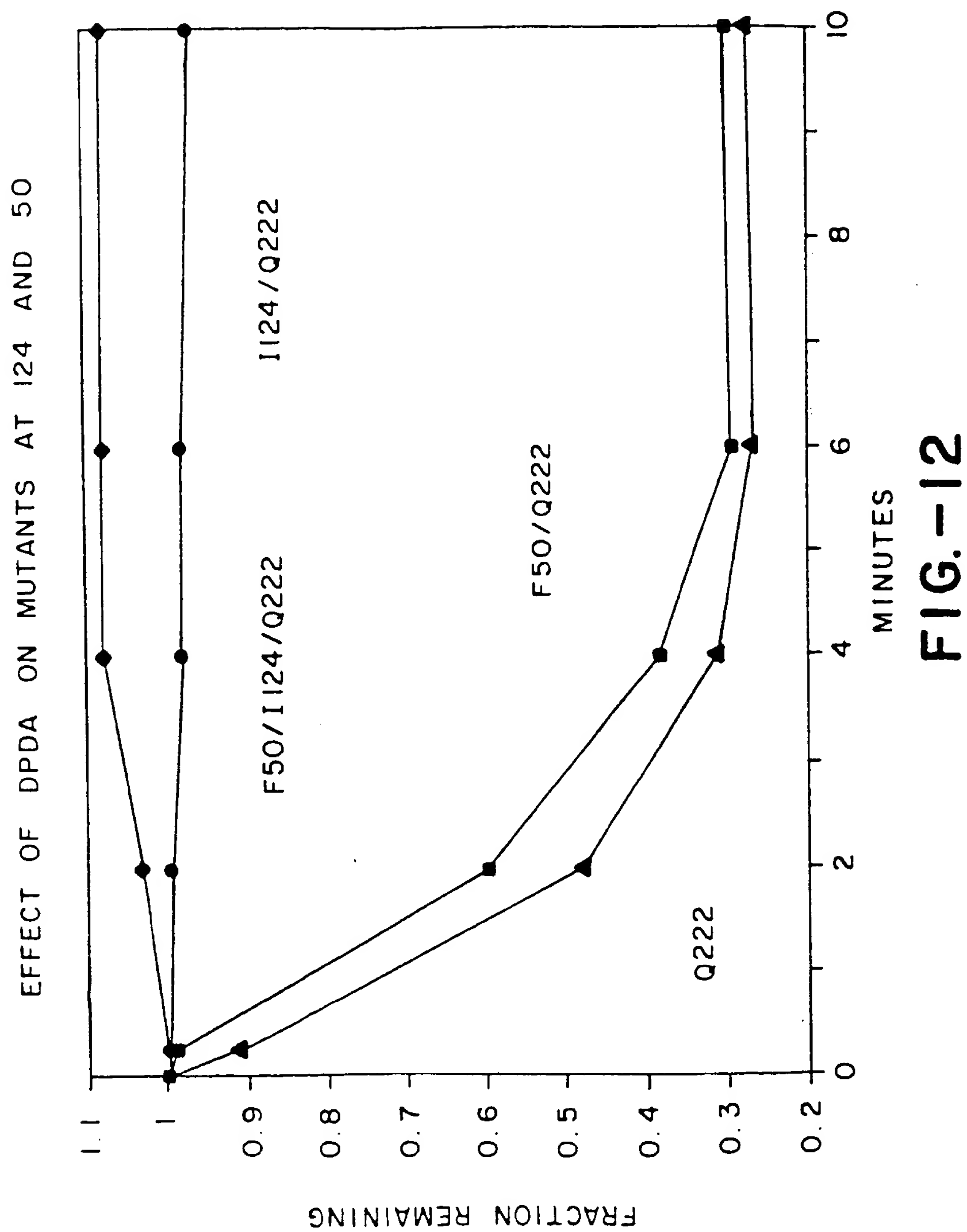


FIG.-12

- Codon: 166
- Wild type amino acid sequence: Thr Ser Gly Ser Ser Thr Val Gly Tyr Pro Gly
1. Wild type DNA sequence:
 5'-ACT TCC GGC AGC TCA AGC ACA GTG GGC TAC CCT GGT-3'
 3'-TGA AGG CCG TCG AGT TCG TGT CAC CCG ATG GGA CCA-5'
 2. pΔ166 DNA sequence:
 * 5'-ACT TCC GGG AGC TCA A-----C CCG GGT-3'
 3'-TGA AGG CCC TCG AGT T-----G GGC CCA-5'
 SacI XmaI
 3. pΔ166 cut with SacI and XmaI:
 * 5'-ACT TCC GGG AGC T
 3'-TGA AGG CCCp pCCG GGT-3'
 CA-5'
 4. Cut pΔ166 ligated with
 duplex DNA cassette pools:
 * 5'-ACT TCC GGG AGC TCA AGC ACA GTG NNN TAC CCG GGT-3'
 3'-TGA AGG CCC TCG AGT TCG TGT CAC NNN ATG GGC CCA-5'
 *** *

MUTAGENESIS PRIMER 37 MER

5' AA GGC ACT TCC GGG AGC TCA ACC CGG GTA AA TAC CCT 3'

FIG.—13

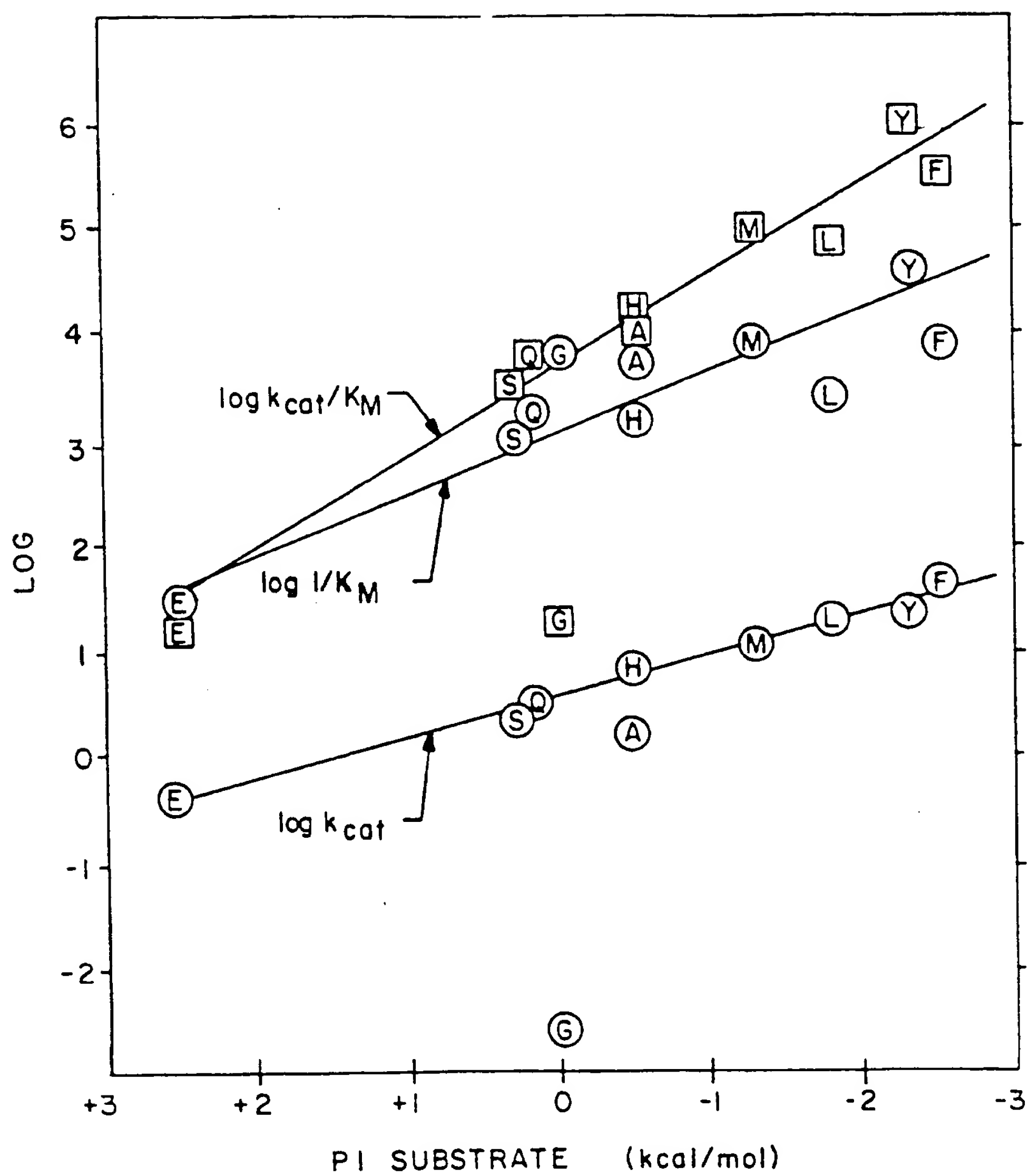
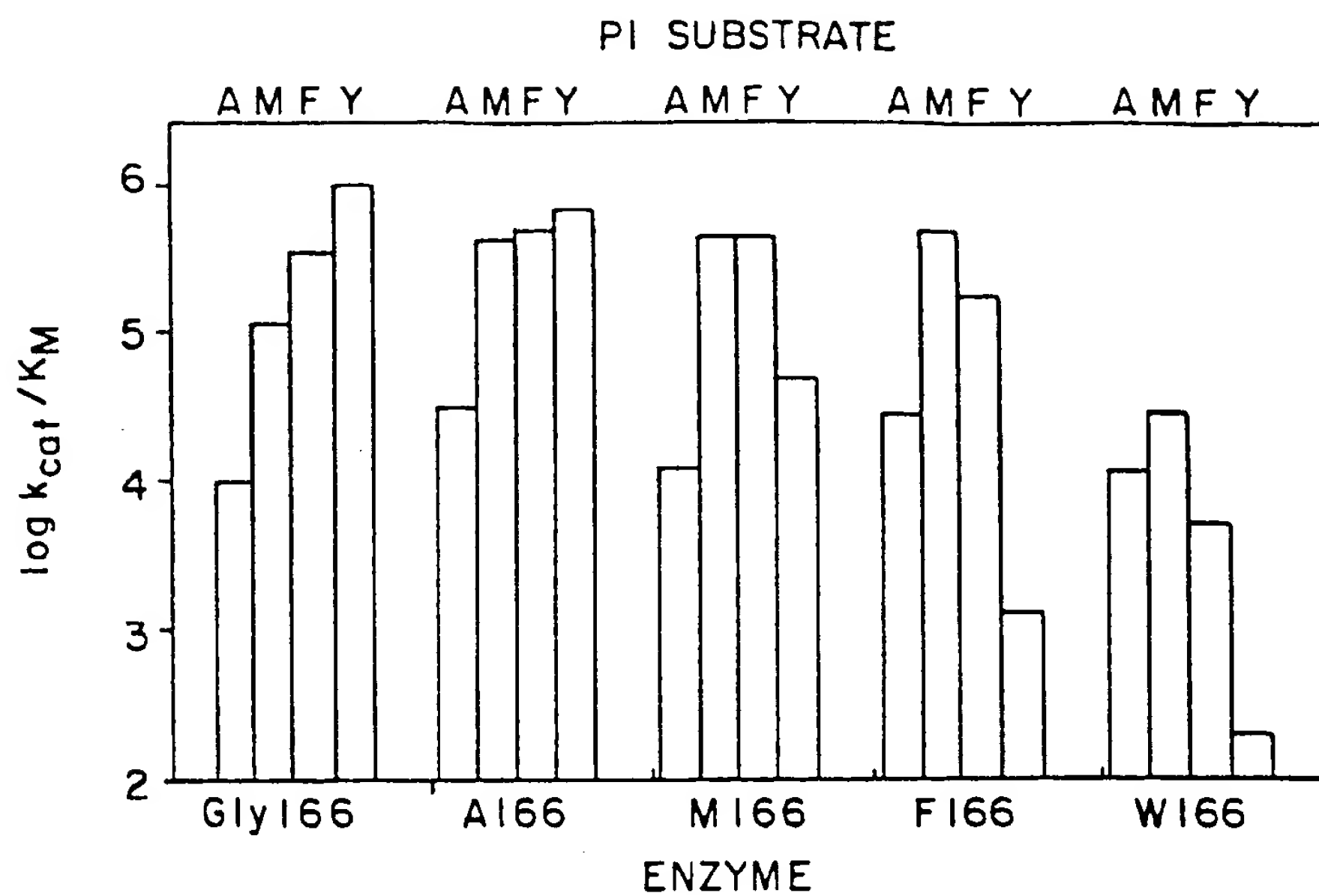
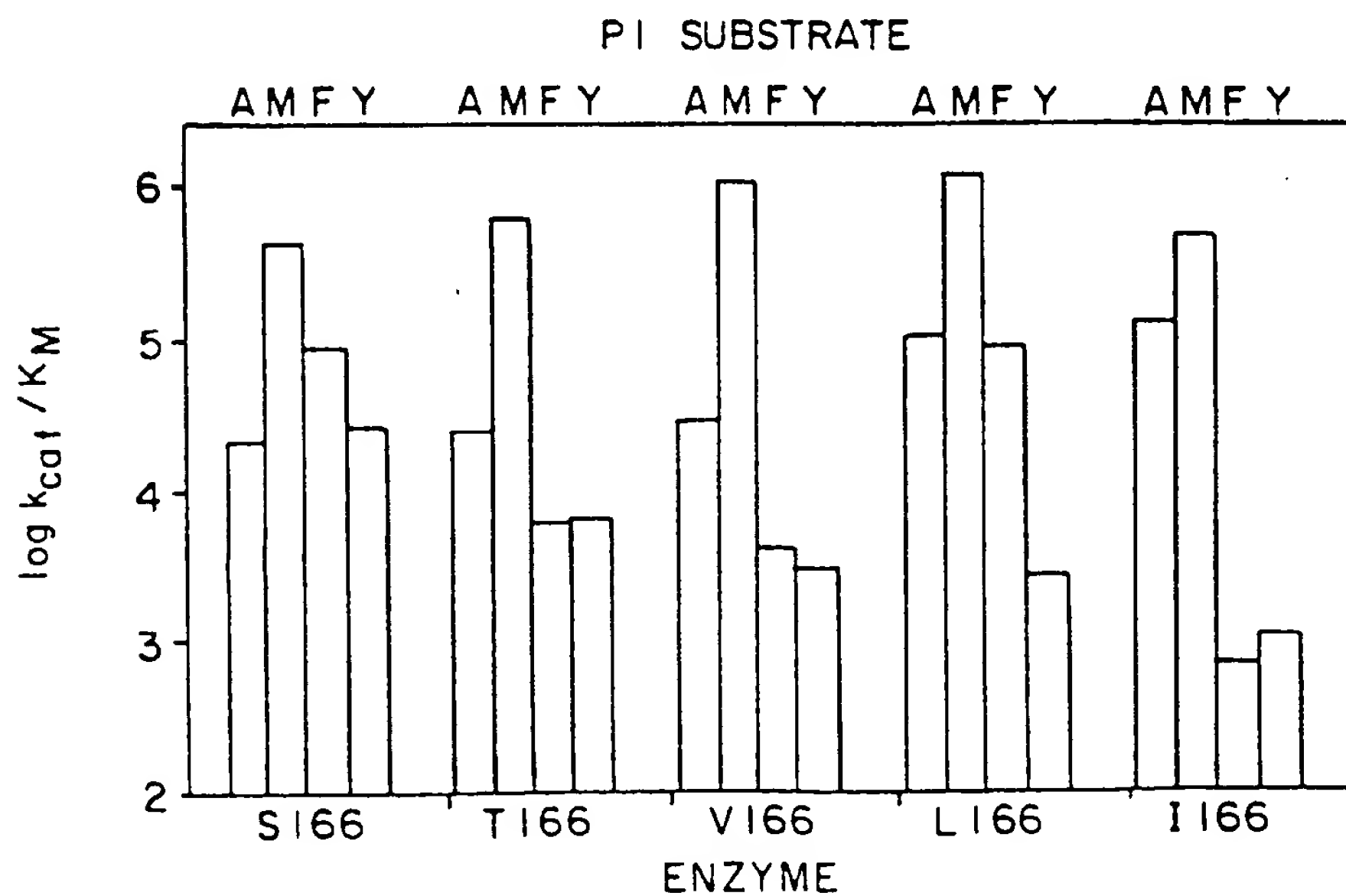


FIG. - 14

**FIG.-15A****FIG.-15B**

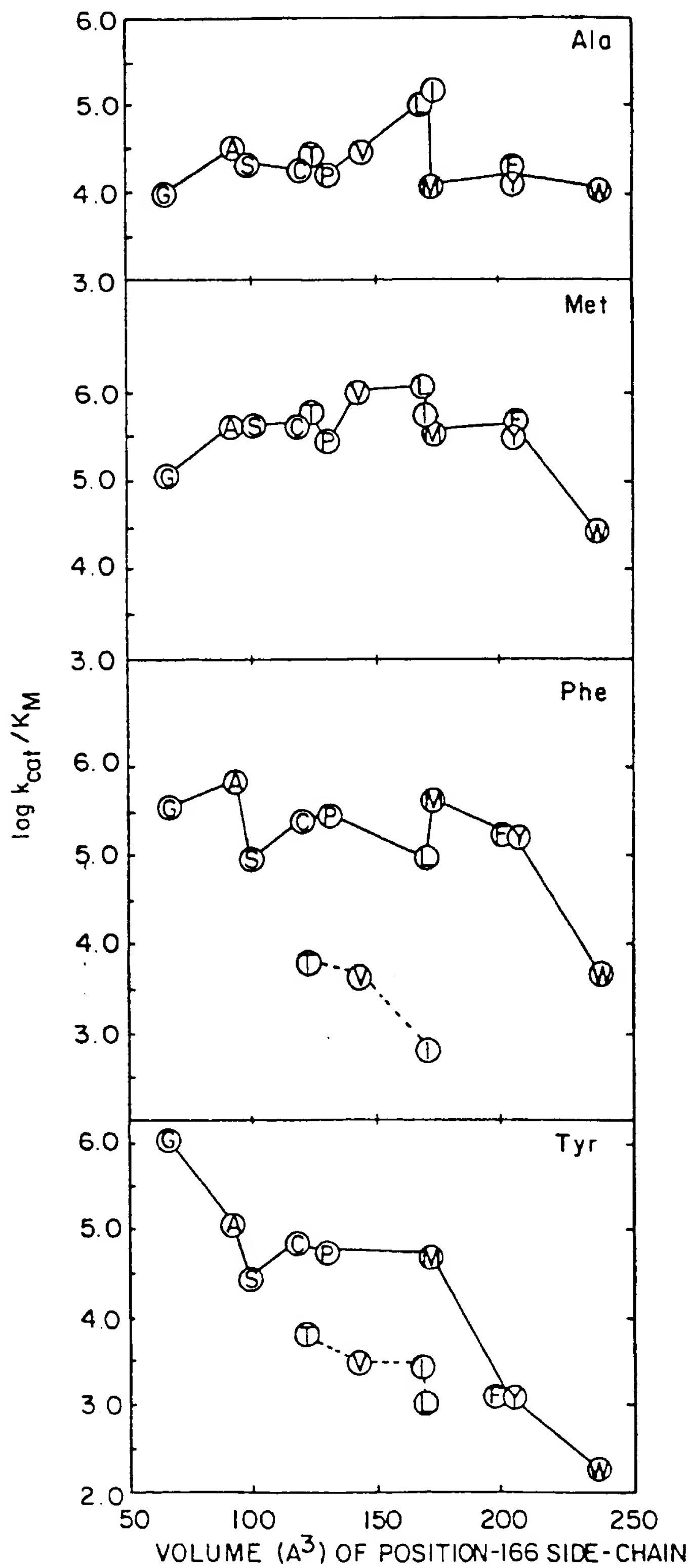


FIG.-16

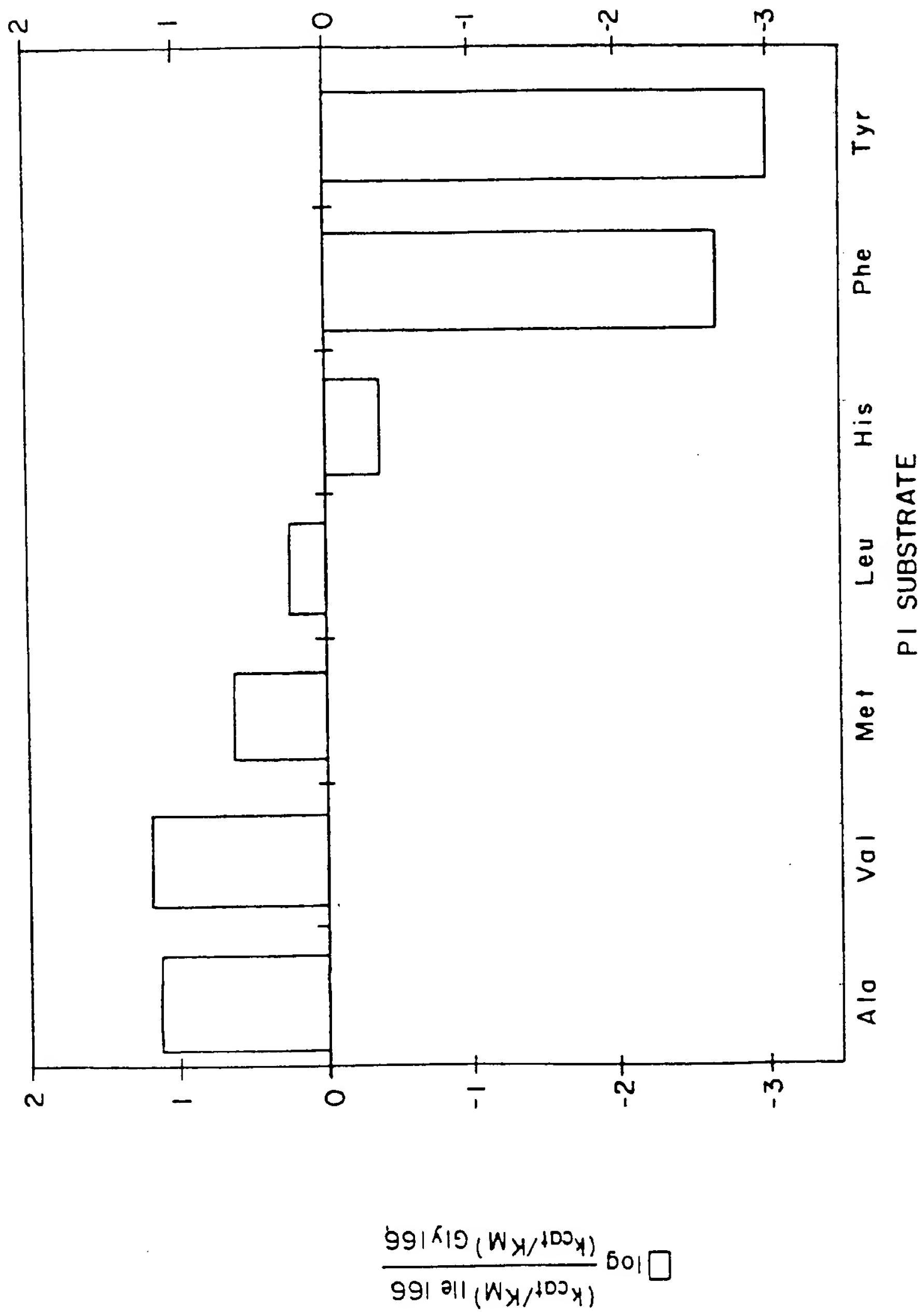


FIG. - 17

GLY-169 CASSETTE MUTAGENESIS

		162													169													173													
		SER SER THR VAL GLY TYR PRO GLY LIS TYR PRO SER																																							
WILD TYPE AMINO ACID SEQUENCE:																																									
1. WILD TYPE DNA SEQUENCE		5'	TCA	AGC	ACA	GTG	GGC	TAC	CCT	GGT	AAA	TAC	CCT	TCT	3'																										
		3'	AGT	TCG	TGT	CAC	CCG	ATG	GGA	CCA	TTT	ATG	GGA	AGA	5'																										
2. P169 DNA SEQUENCE		5'	TCA	AGC	ACA	GTC	GGG	TAC	CCT	-----GA	TAT	CCT	TCT	3'																											
		3'	AGT	TCG	TGT	CAC	CCC	ATG	GGA	CT	ATA	GGA	AGA	5'																											
			KPN1													ECORV																									
3. P169 CUT WITH KPN1 AND ECORV:		5'	TAC	AGC	ACA	GTC	GGG	TAC														PAT	CCT	TCT	3'																
		3'	AGT	TCG	TGT	CAC	CCP														TA	GGA	AGA	5'																	
4. CUT P169 LIGATED WITH OLIGONUCLEOTIDE POOLS		5'	TAC	AGC	ACA	GTG	GGG	TAC	CCT	NNN	AAA	TAT	CCT	TGT	3'																										
		3'	AGT	TCG	TGT	CAC	CCC	ATG	GGA	NNN	TTT	ATA	GGA	AGA	5'																										
MUTAGENESIS PRIMER FOR P169		5'	AAG	CAC	AGT	GGG	GTA	CCC	TGA	TAT	CCT	TCT	GTC	A	3'																										

FIG.—18

1. Codon number: 100 104 105 108
2. Wild type amino acid sequence: Gly-Ser-Gly-Gln-Tyr-Ser-Trp-Ile-Ile-
3. Wild type DNA sequence: 5'-GGT-TCC-GGC-CAA-TAC-AGC-TGG-ATC-ATT-3'
Pvu II
4. Primer for *Hind* III
insertion at 104:

 5'-GGT-TCC-GGC-CAA-GCTT-AGC-TGG-ATC-ATT-3'
 Hind III
5. Primers for 104 mutants:

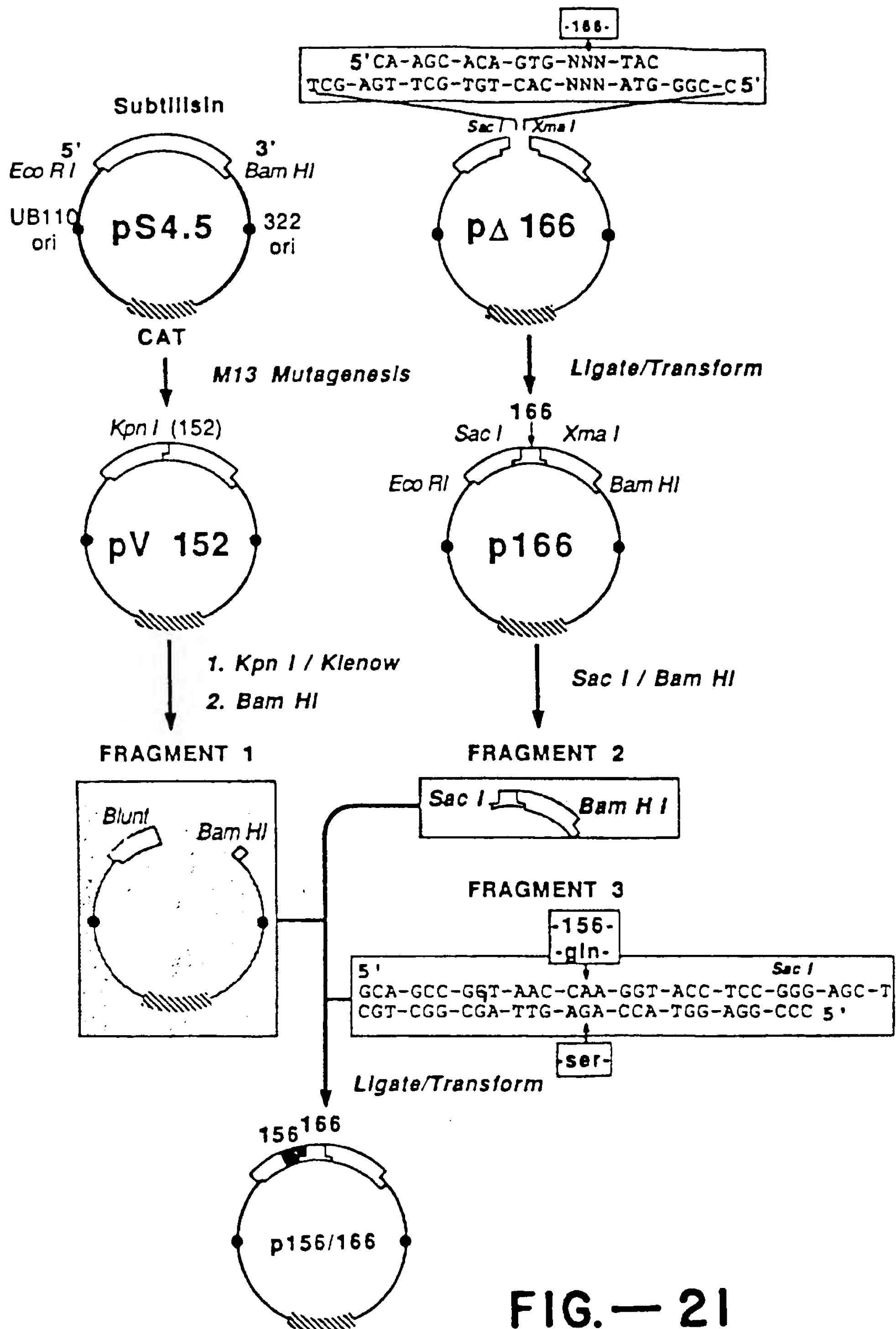
**
 5'----T-TCC-GCC-CAA-NNN-AGC-TGG-ATC-----3'
6. Mutants made: A, M, L, S, AND H104

FIG.—19

1. Codon number: 148 150 152 155
2. Wild type amino acid sequence: Val-Val-Val-Ala-Ala-Ala-Gly-Asn-Glu
3. Wild type DNA sequence: 5'-GTA-GTC-GTT-GCG-GCA-GCC-GGT-AAC-GAA-3'
4. V152/P153
 5'-GTA-GTC-GTT-GCG-GTA-CCC-GGT-AAC-GAA-3'
 * * *
 Kpn I
5. S152:

 5'-GTA-GTC-GTT-GCG-AGC-GCC-GGT-AAC-GAA-3'
6. G152:
 **
 5'-GTA-GTC-GTT-GCG-GGC-GCC-GGT-AAC-GAA-3'

FIG.-20



1. Codon number: 211 215 217 220
2. Wild type amino acid sequence: Gly-Asn-Lys-Tyr-Gly-Ala-Tyr-Asn-Gly-Thr-Ser-Met-Ala
3. Wild type DNA sequence: 5'-GGA-AAC-AAA-TAC-GGG-GCG-TAC-AAC-GGT-ACG-TCA-ATG-GCA
CCT-TTG-TTT-ATG-CCC-CGC-ATG-TTG-CCA-TGC-AGT-TAC-CGT-5'
4. pΔ217
5'-GGA-AAC-AAA-TAC-GGC-GCC-TAC-----GG-ATA-TCA-ATG-GCA
CCT-TTG-TTT-ATG-CCG-CGG-ATG-----CC-TAT-AGT-TAC-CGT-5'
Nar I Eco RV
5. pΔ217 cut with Nar I and Eco RI
5'-GGA-AAC-AAA-TAC-GG*
CCT-TTG-TTT-ATG-CCG-Gp
* PA-TCA-ATG-GCA
T-AGT-TAC-CGT-5'
6. Cut pΔ217 ligated with cassettes:
5'-GGA-AAC-AAA-TAC-GGC-GCG-NNN-AAC-GGT-ACA-TCA-ATG-GCA
CCT-TTG-TTT-ATG-CCG-CGC-NNN-TTG-CCA-TGT-AGT-TAC-CGT-5'

7. Mutagenesis primer for pΔ217:
5'-GA-AAC-AAA-TAC-GGC-GCC-TAC-GGA-TAT-CAA-TGG-CAT-3'
* * *
8. Mutants made: All 19 at 217

FIG.-22

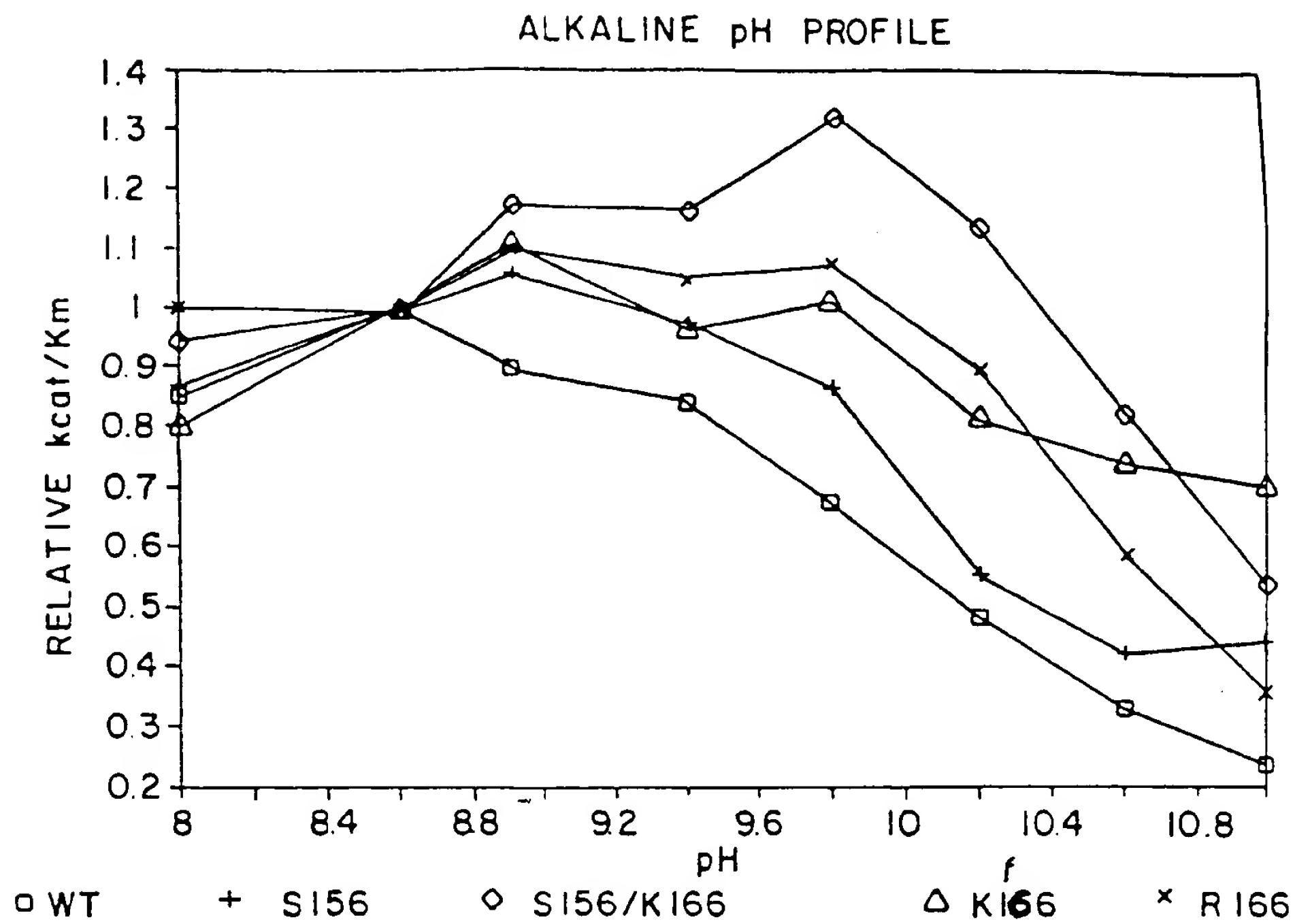


FIG. - 23A

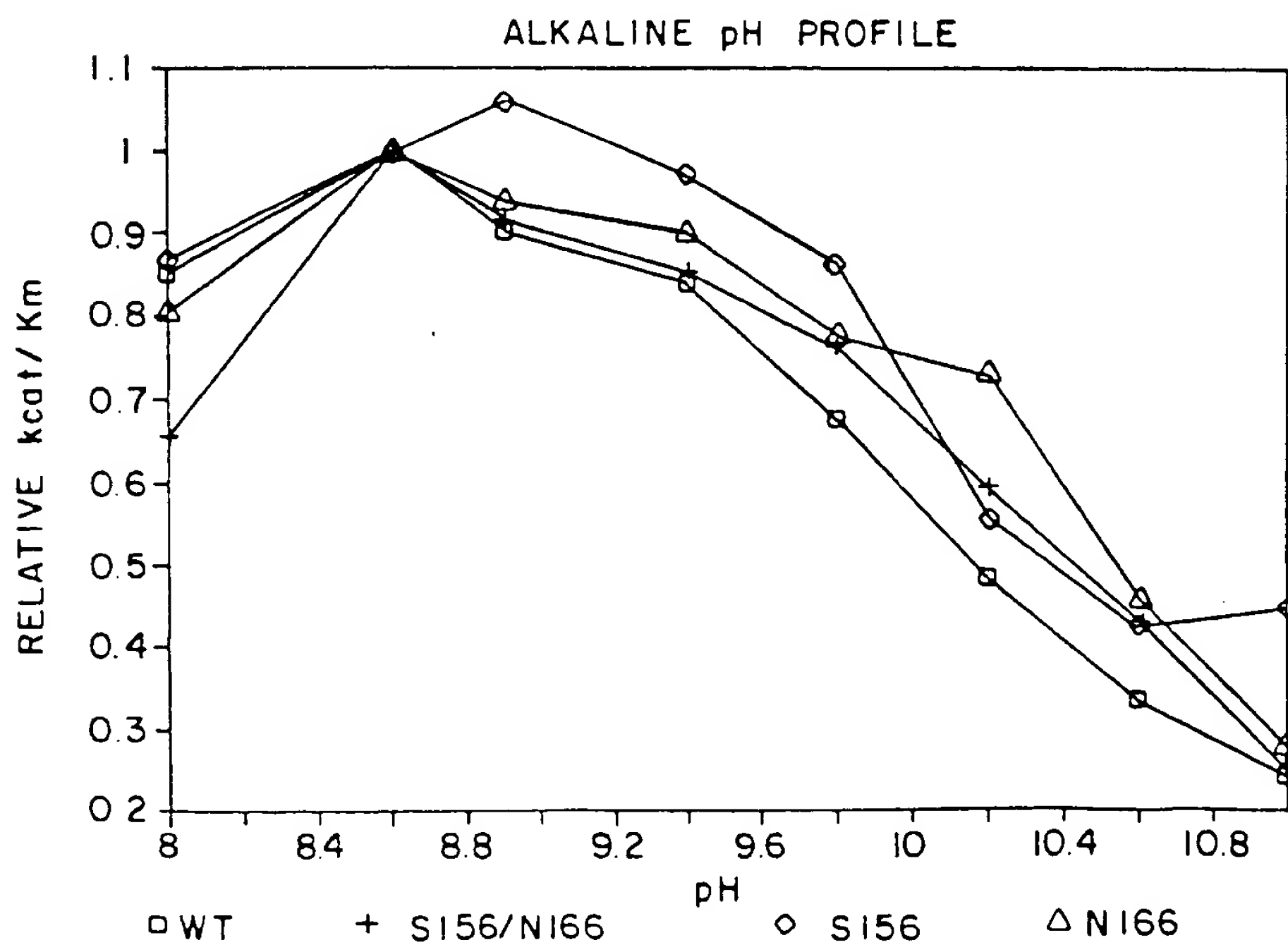


FIG. - 23B

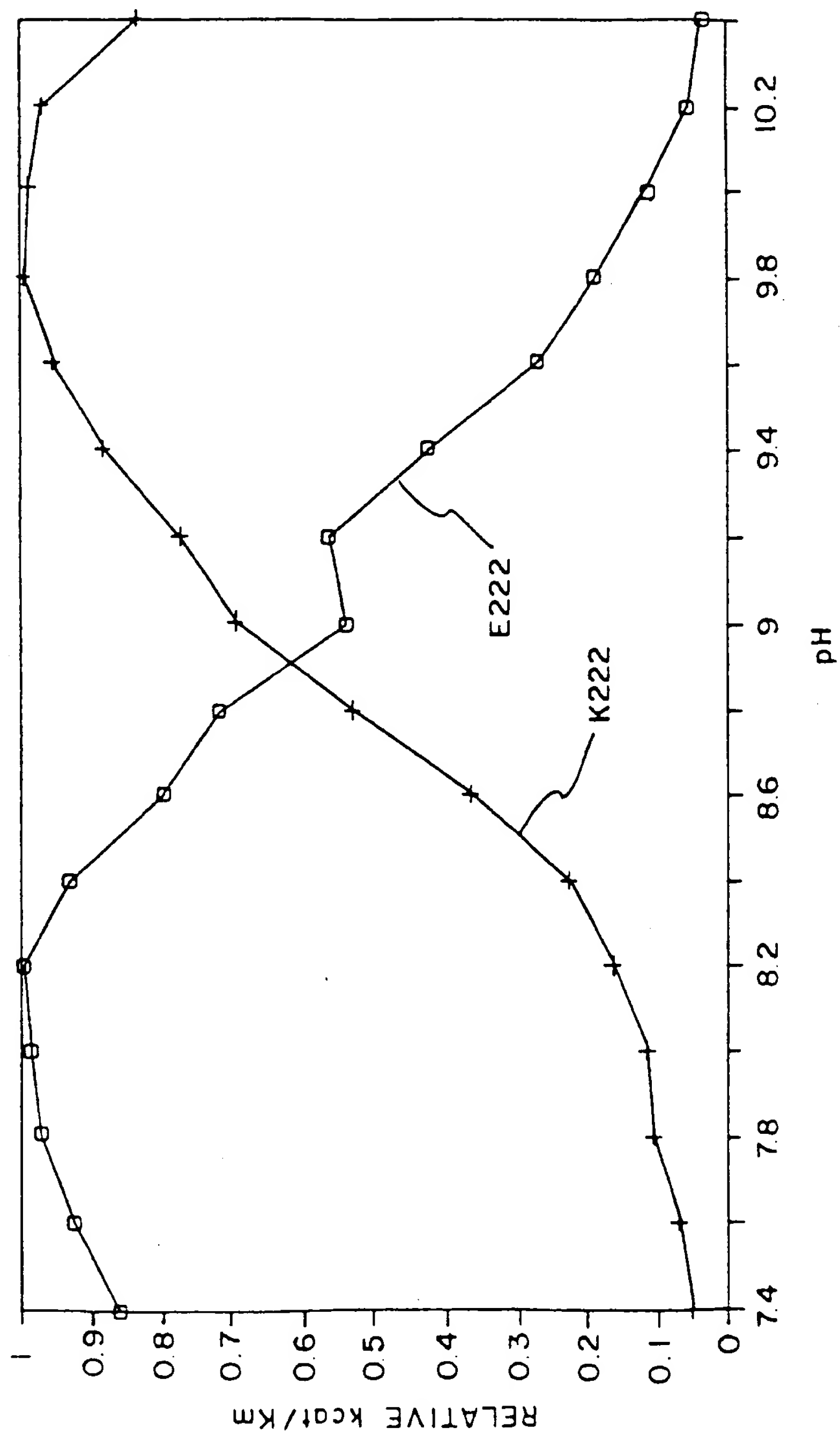


FIG. - 24

1. Codon number: 91 95 100
2. Wild type amino acid sequence: Tyr-Ala-Val-Lys-Val-Leu-Gly-Ala-Asp-Gly-Ser
3. Wild type DNA sequence: 5'-TAC-GCT-GTA-AAA-GTT-CTC-GGT-GCT-GAC-GGT-TCC
ATG-CGA-CAT-TTT-CAA-GAG-CCA-CTG-CCA-AGG-5'
4. pΔ95: 5'-TAC-GCG-T-CTC-GCT-GCA-GAC-GGT-TCC
ATG-CGC-A-GAG-CTG-CCA-AGG-5'
MuI
5. pΔ95 cut with *MuI* and *Pst* I 5'-TA-ATG-CGCP
A-CGT-CTG-CCA-AGG-5'
6. Cut pΔ95 ligated with cassettes: 5'-TAC-GCG-GTA-AAA-GTT-CTC-GGT-GCA-GAC-GGT-TCC
ATG-CGC-CAT-TTT-CAA-GAG-CCA-CTG-CCA-AGG-5'
7. Mutagenesis primer for pΔ95: 5'-CA-TCA-CTT-TAC-GCG-T-CTC-GCT-GCA-GAC-GGT-TCC
8. Mutants made: C94, C95, D96

FIG.-25

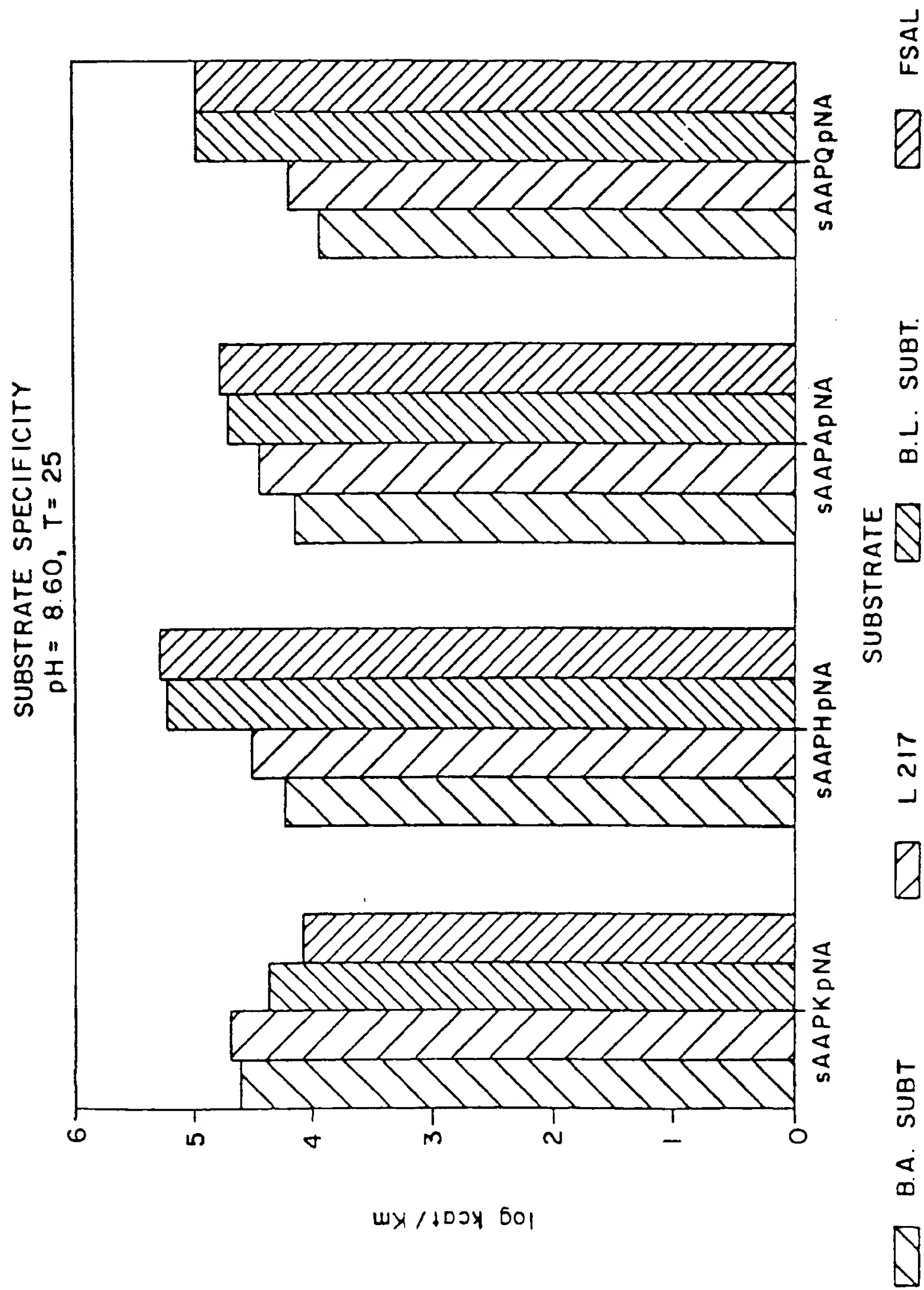


FIG.-26

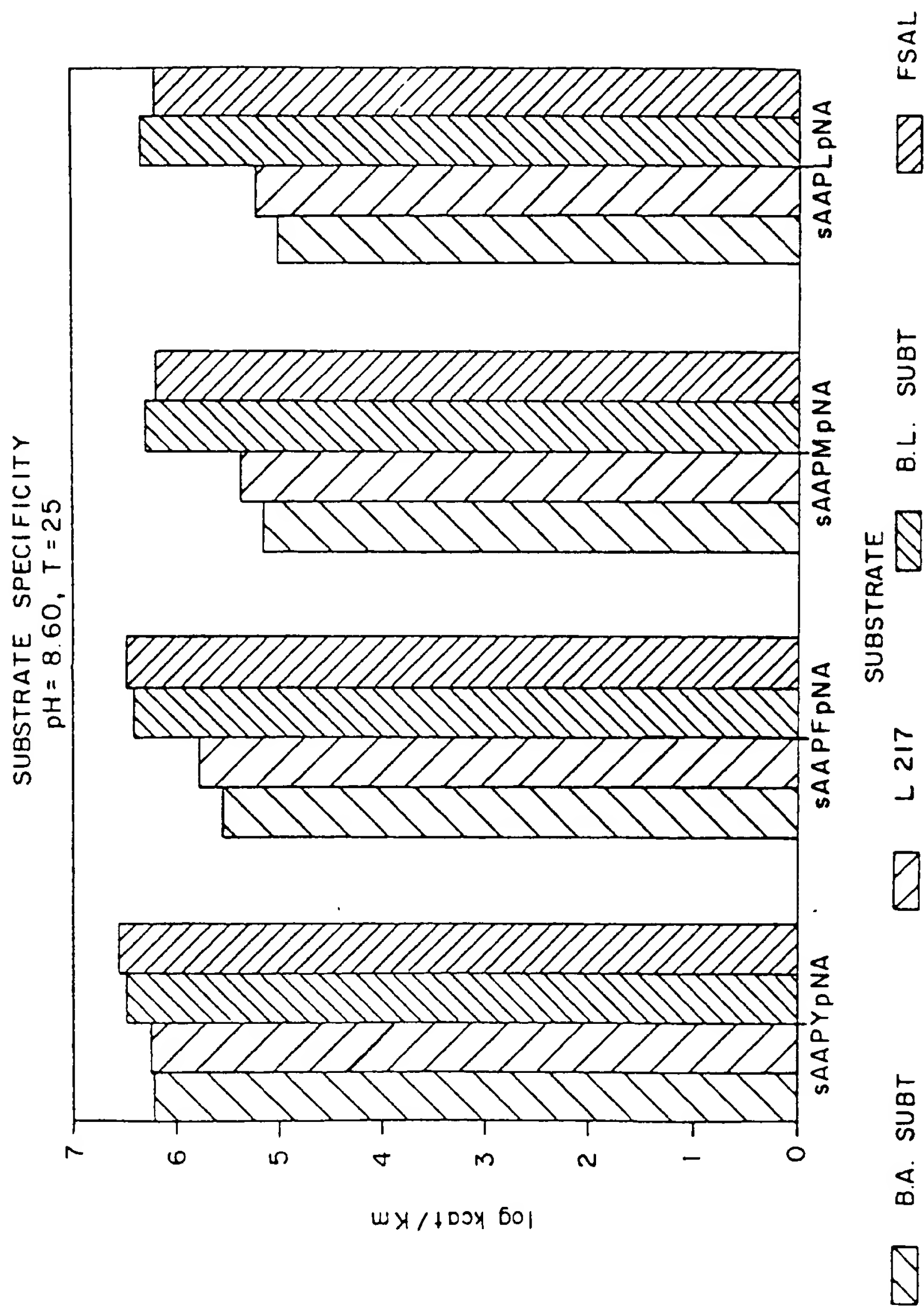


FIG.-27

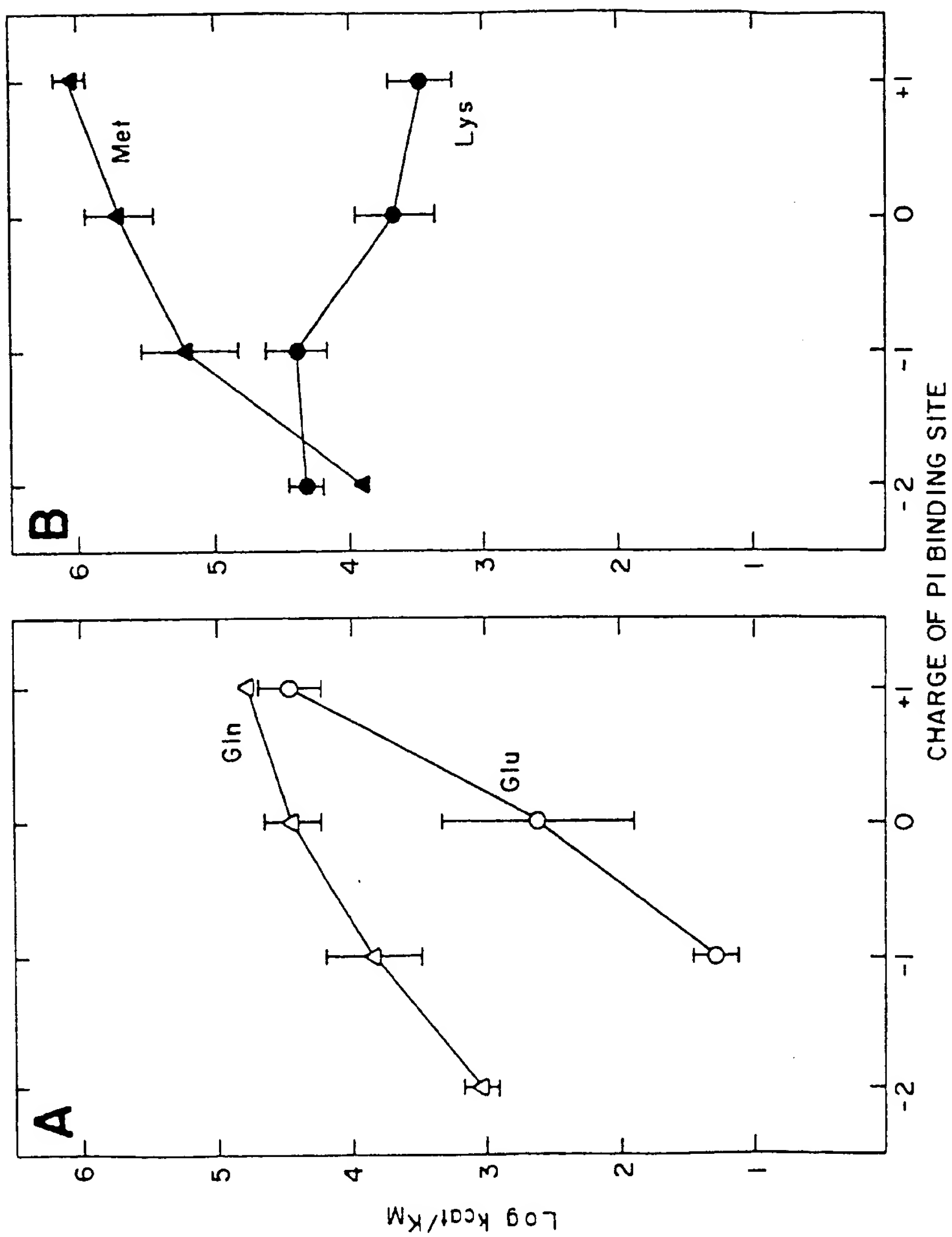


FIG.-28

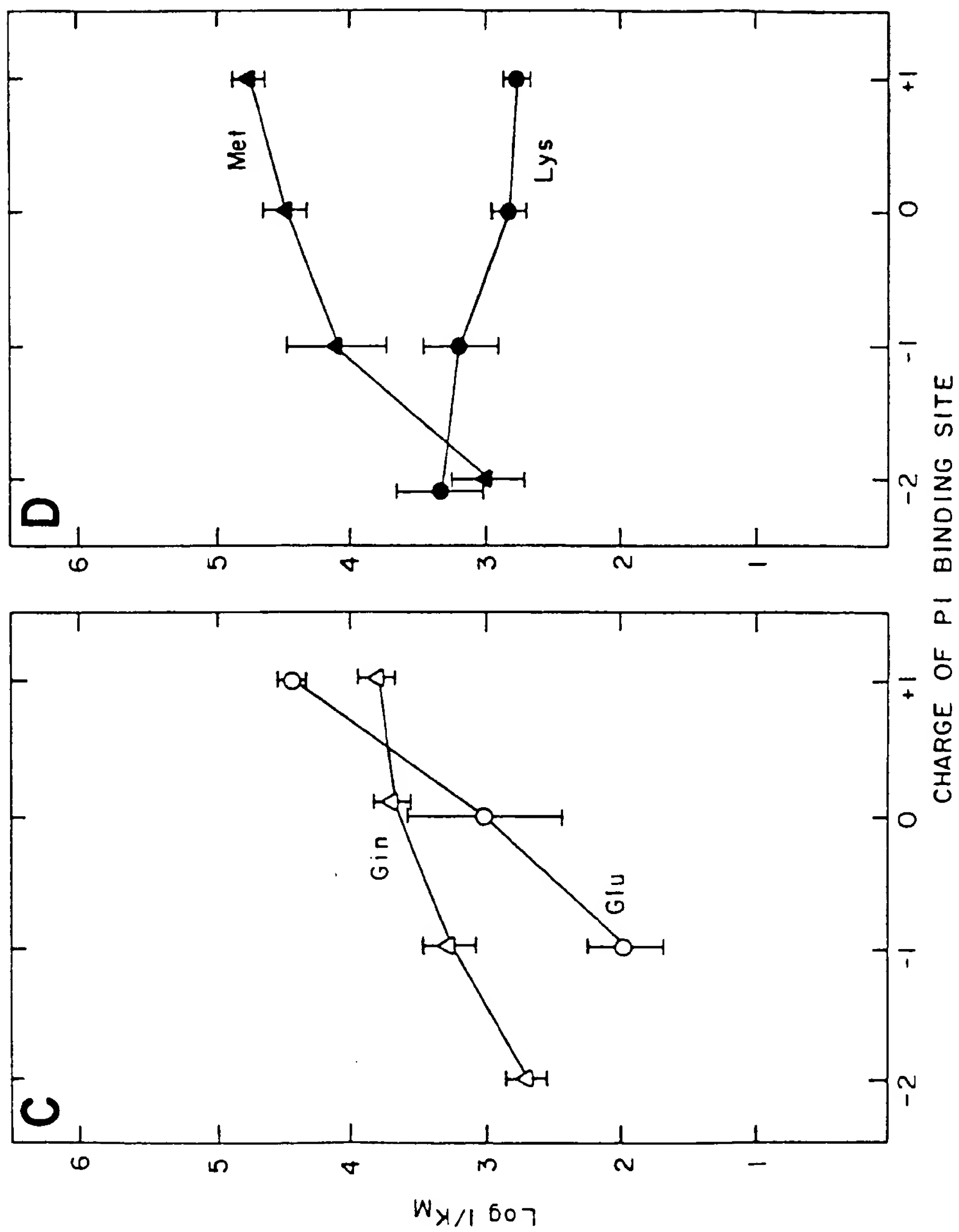


FIG.-28

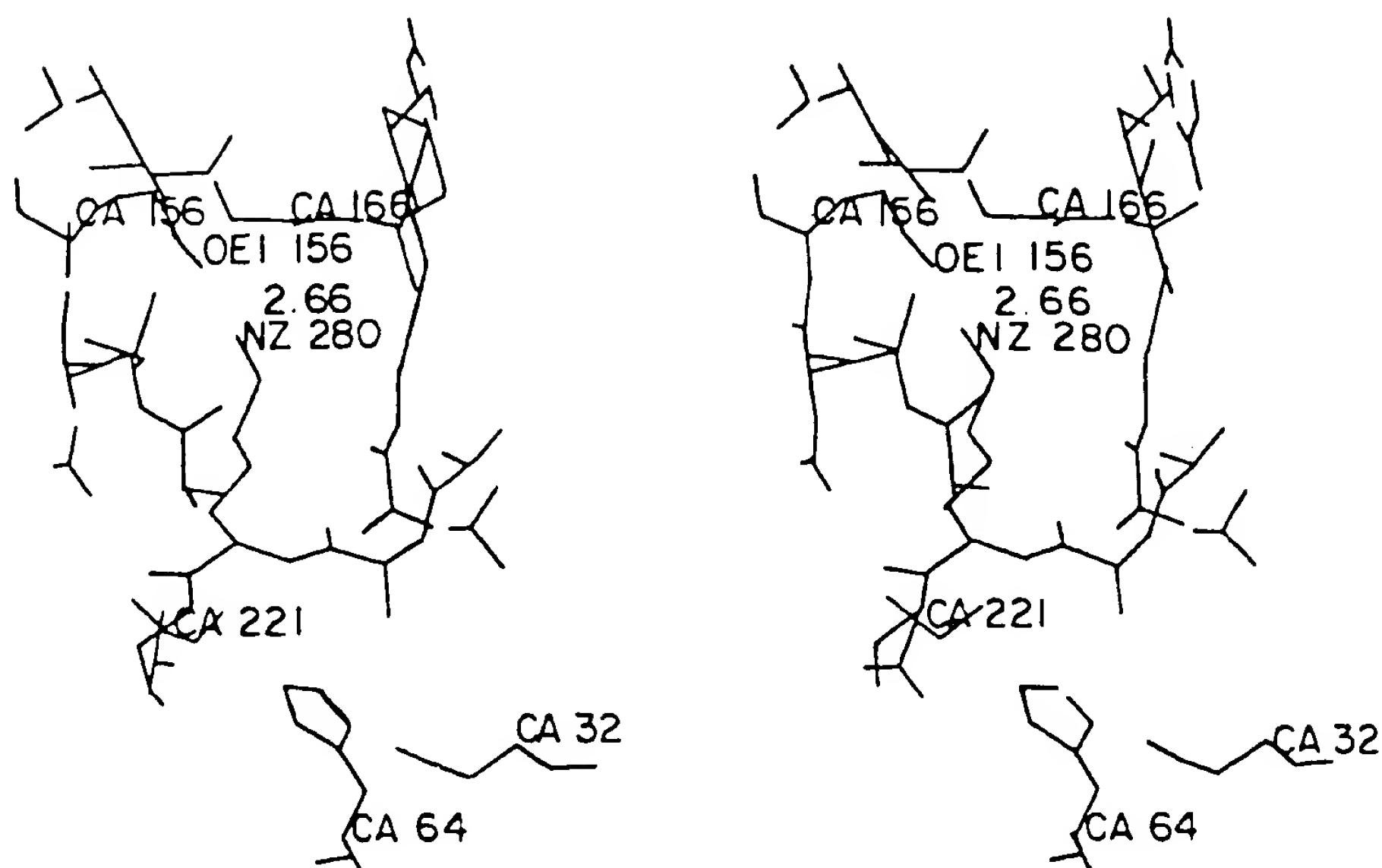


FIG. — 29A

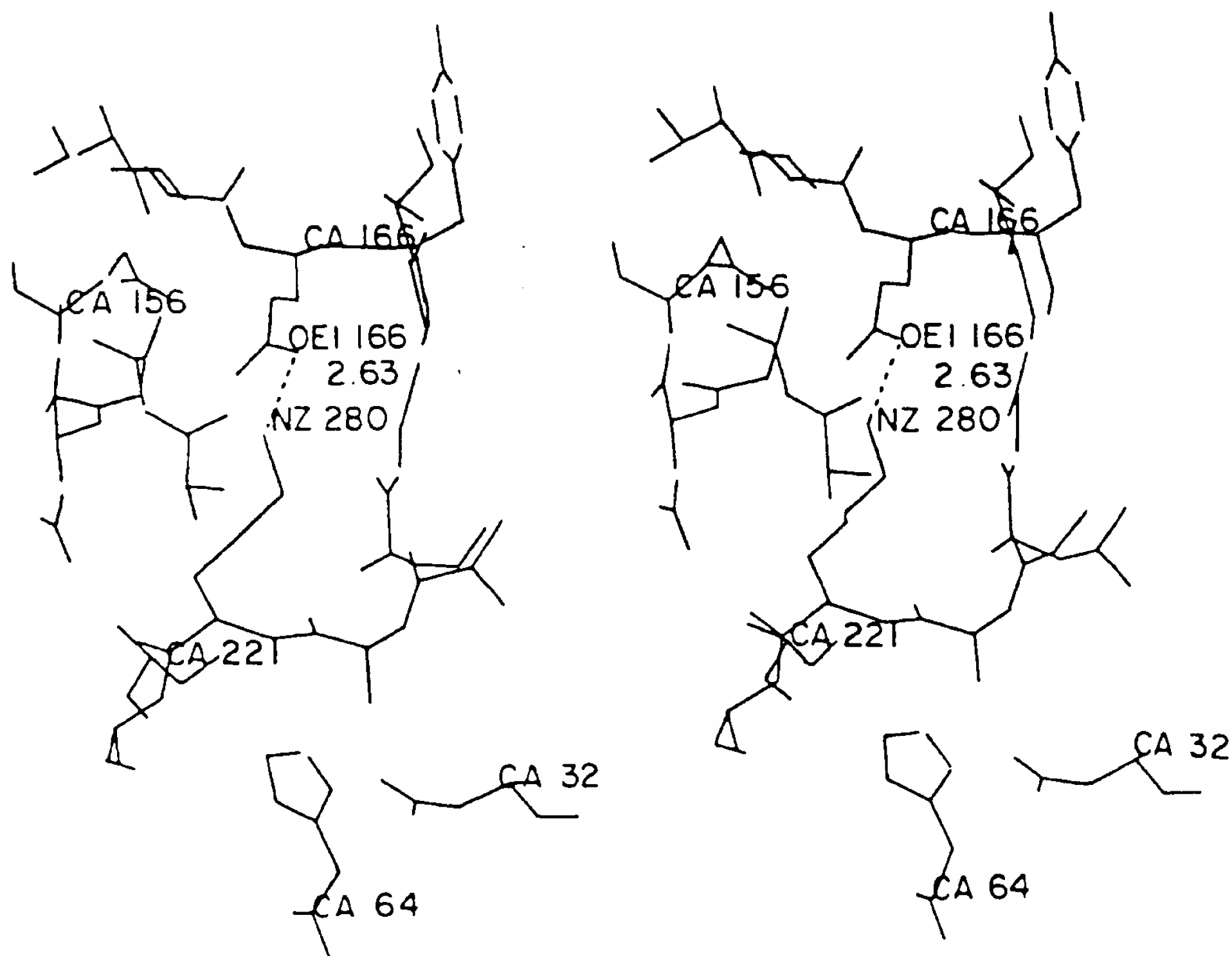


FIG. — 29B

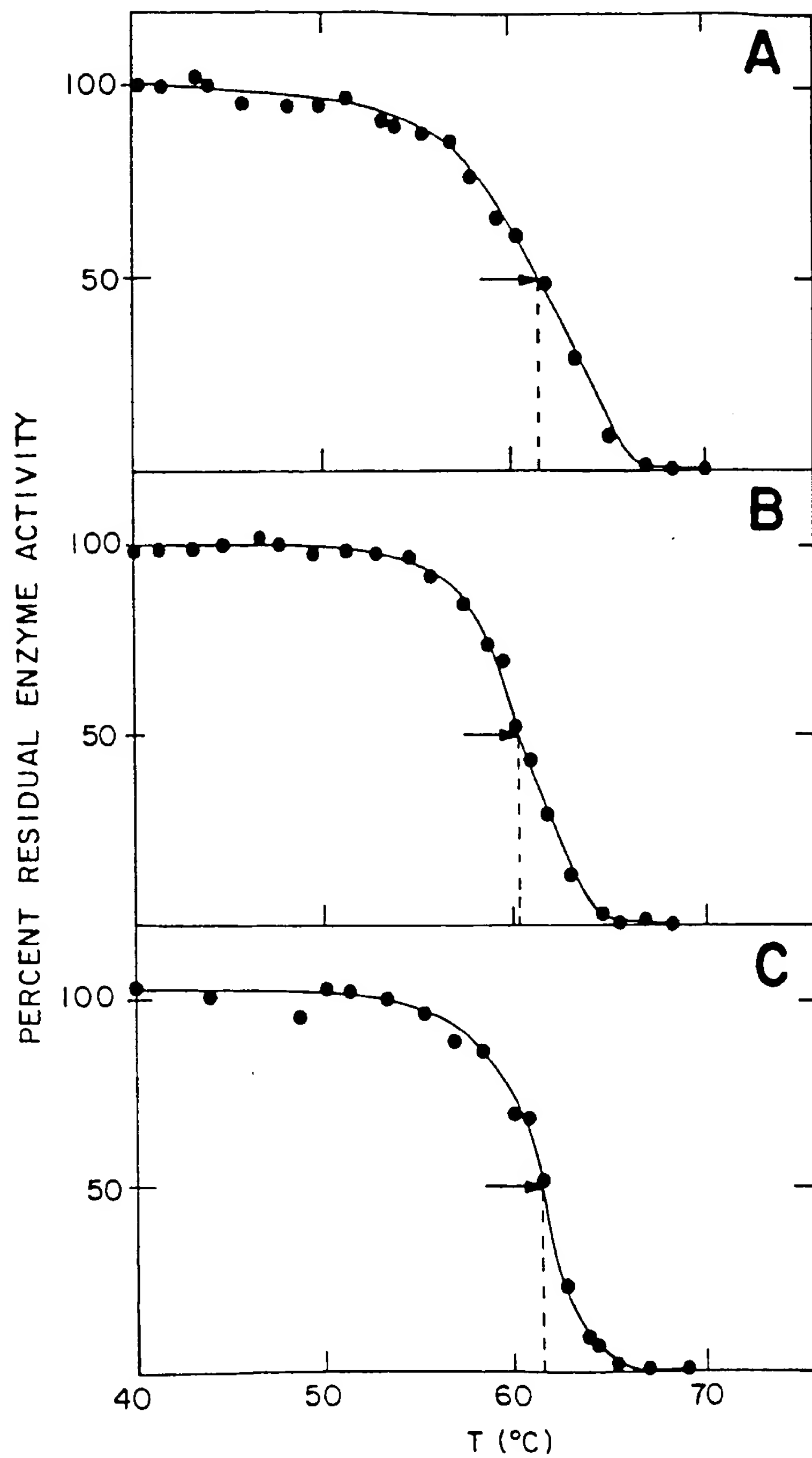


FIG.—30

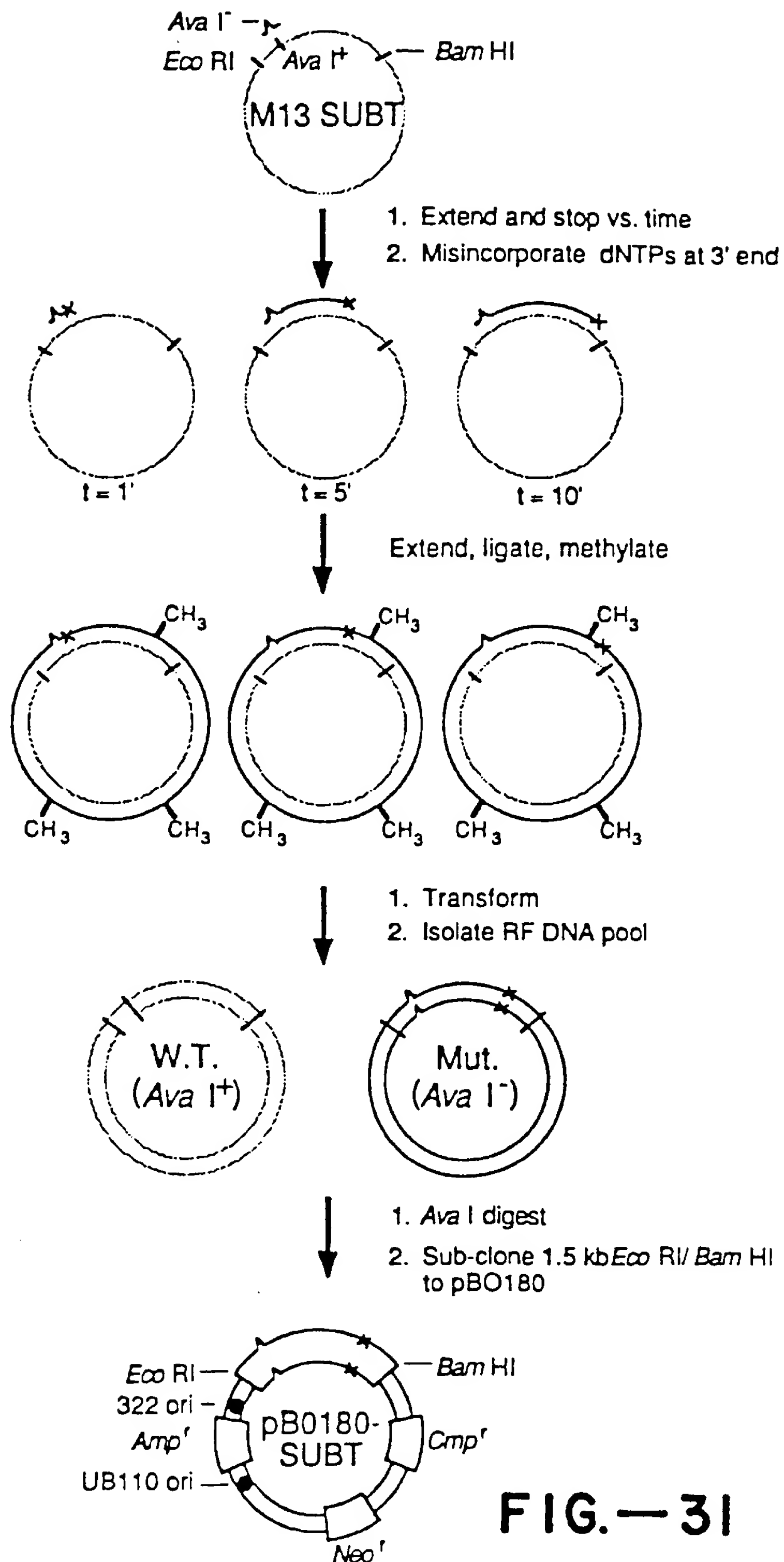


FIG.—31

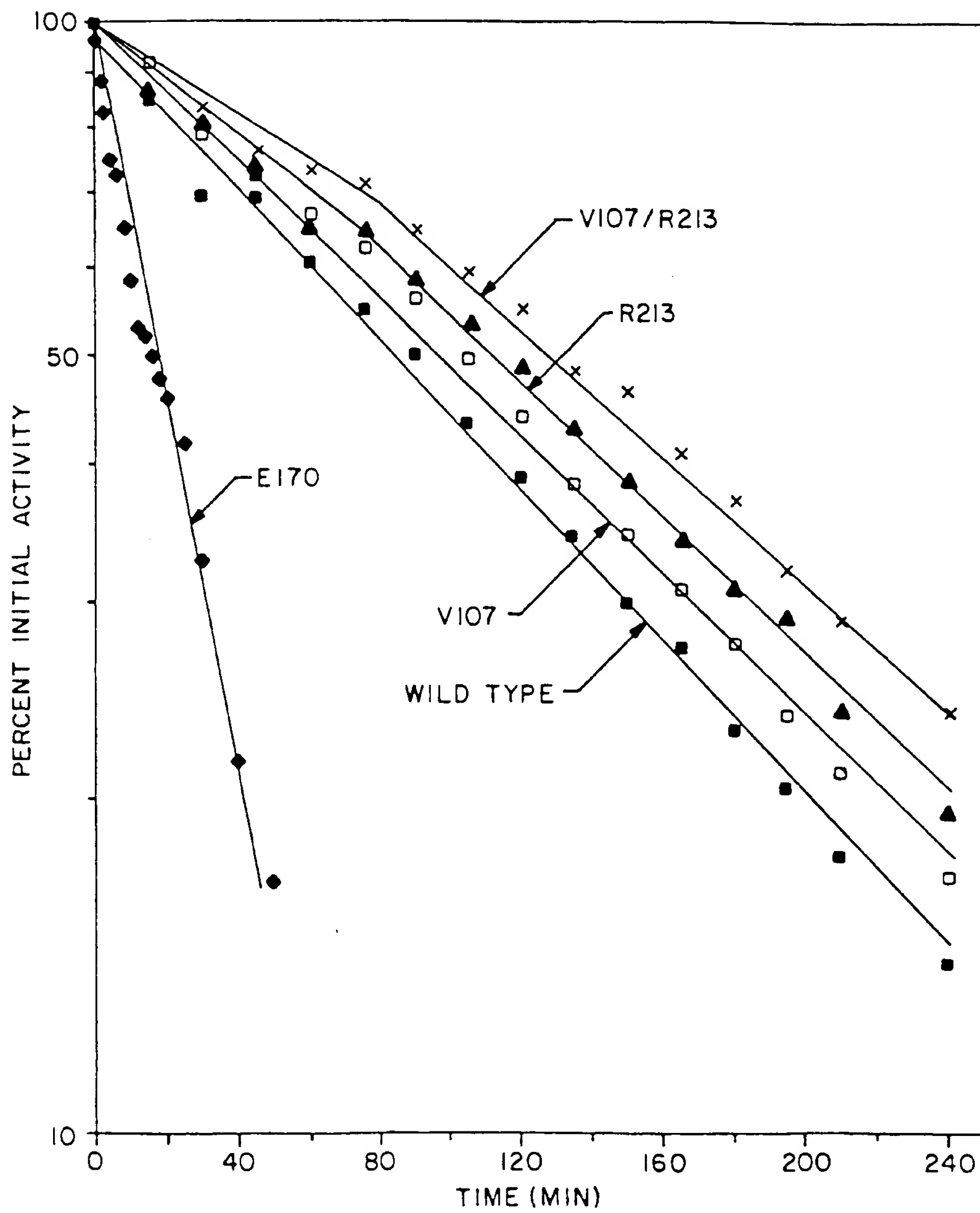


FIG.-32

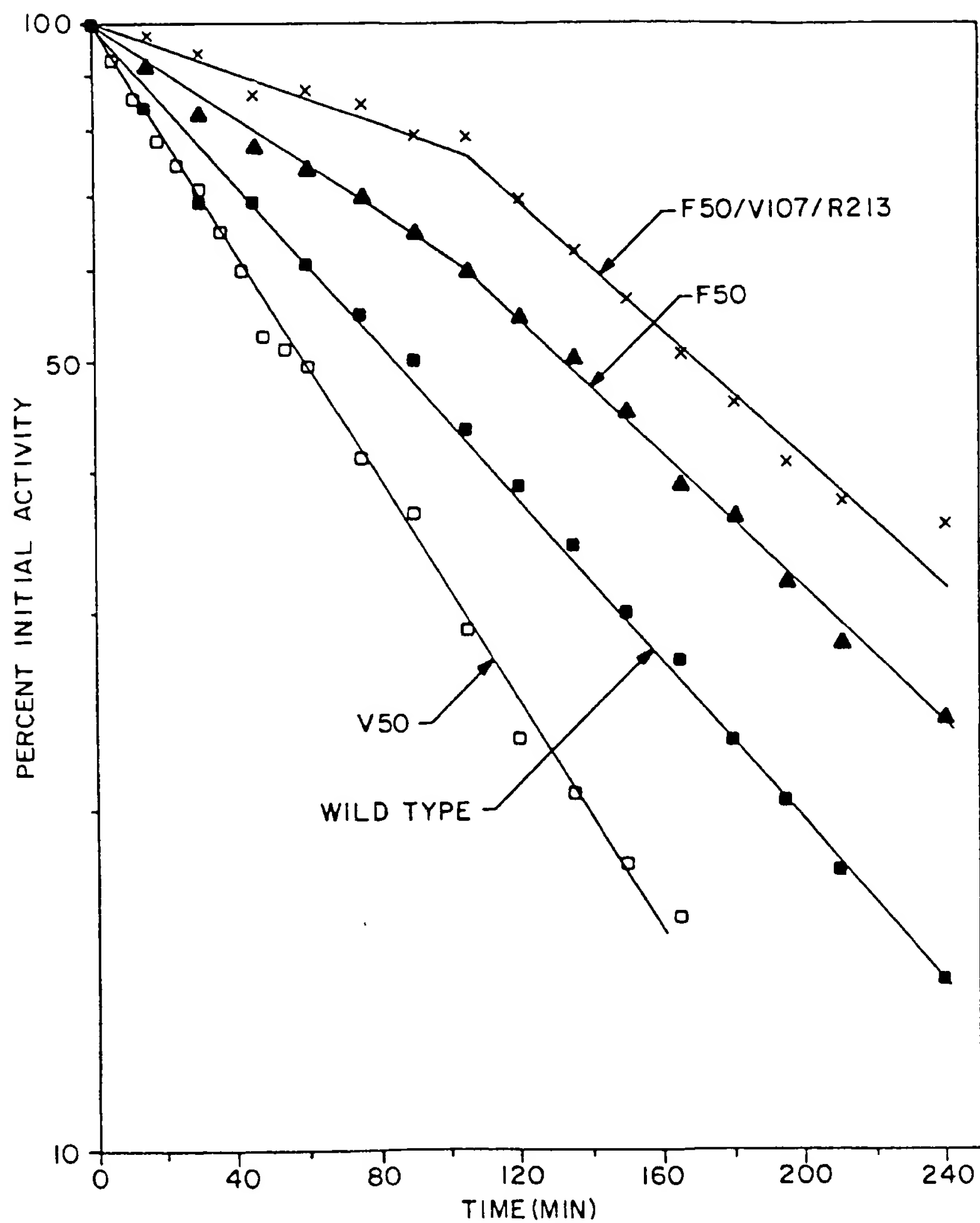


FIG.-33

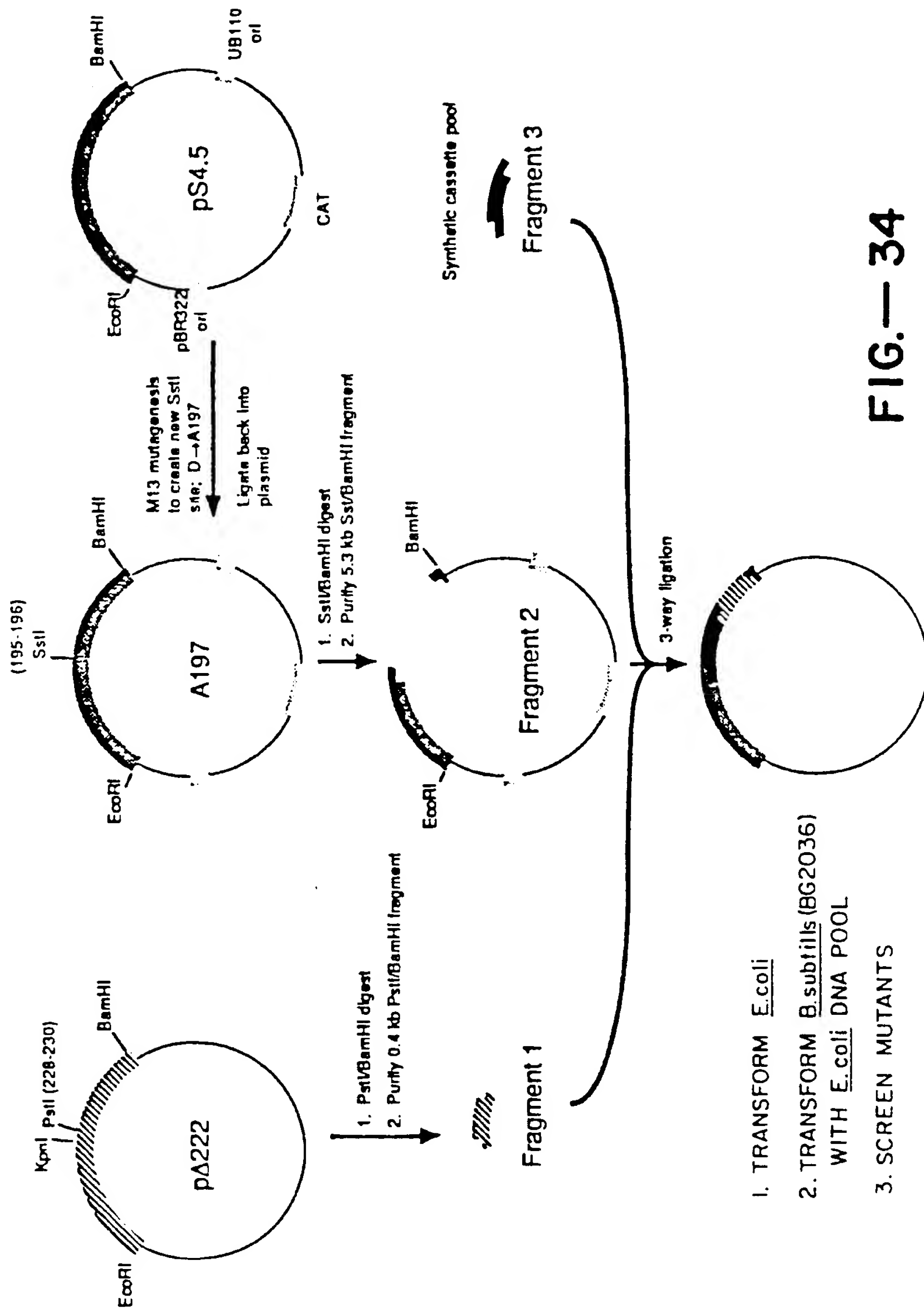


FIG.— 34

	195	200	206
W.T A.A.:	Glu	Leu Asp Val Met Ala Pro Gly Val Ser Ile	Gln
W.T. DNA:	GAG CTT GAT GTC ATG GCA CCT GGC GTA TCT ATC CAA		
	CTC GAA CTA CAG TAC CGT GGA CCG CAT AGA TAG GTT		
pΔ222DNA:	GAG CTT GAT GTC ATG GCA CCT GGC GTA TCT ATC CAA		
	CTC GAA CTA CAG TAC CGT GGA CCG CAT AGA TAG GTT		
A197 DNA:	<u>GAG CTC</u> [*] GCA GTC ATG GCA CCT GGC GTA TCT ATC CAA		
	CTC GAG CGT CAG TAC CGT GGA CCG CAT AGA TAG GTT		
	<i>SstI</i>		
Fragments from pΔ222 and A197 cut w/ <i>PstI</i> , <i>SstI</i> :	GAG-CT		
	Cp		
	[*]		
pΔ222, A197 cut & ligated w/ oligodeoxy- nucleotide pools:	<u>GAG CTC</u> GAT GTC ATG GCA CCT GGC GTA TCT ATC CAA		
	CTC GAG CTA CAG TAC CGT GGA CCG CAT AGA TAG GTT		
	<i>SstI</i>		
	207	210	218
W.T A.A.:	Ser Thr Leu Pro Gly Asn Lys Tyr Gly Ala Tyr Asn		
W.T. DNA:	AGC ACG CTT CCT GGA AAC AAA TAC GGG GCG TAC AAC		
	TCG TGC GAA GGA CCT TTG TTT ATG CCC CGC ATG TTG		
pΔ222DNA:	AGC ACG CTT CCT GGA AAC AAA TAC GGG GCG TAC AAC		
	TCG TGC GAA GGA CCT TTG TTT ATG CCC CGC ATG TTG		
A197 DNA:	AGC ACG CTT CCT GGA AAC AAA TAC GGG GCG TAC AAC		
	TCG TGC GAA GGA CCT TTG TTT ATG CCC CGC ATG TTG		
Fragments from pΔ222 and A197 cut w/ <i>PstI</i> , <i>SstI</i> :	[*]		
	<u>AGC ACG CTT CCC GGG</u> AAC AAA TAC GGG GCG TAC AAC		
	TCG TGC GAA GGG CCC TTG TTT ATG CCC CGC ATG TTG		
	<i>SmaI</i>		
	219	220	230
W.T A.A.:	Gly Thr Ser Met Ala Ser Pro His Val Ala Gly Ala		
W.T. DNA:	GGT ACG TCA ATG GCA TCT CCG CAC GTT GCC GGA GCG-3'		
	CCA TGG AGT TAC CGT AGA GGC GTG CAA CGG CCT CGC-5'		
pΔ222DNA:	<u>GGT ACG</u> TCA-----CG CAC <u>GCT GCA</u> GGA GCG-3'		
	CCA TGG AGT-----GC GTG CGA CGT CCT CGC-5'		
	<i>KpnI</i>	<i>PstI</i>	
A197 DNA:	GGT ACG TCA ATG GCA TCT CCG CAC GTT GCC GGA GCG-3'		
	CCA TGG AGT TAC CGT AGA GGC GTG CAA GTG CCT CGC-5'		
Fragments from pΔ222 and A197 cut w/ <i>PstI</i> , <i>SstI</i> :			pGGA GCG-3'
			A CGT CCT CGC-5'
	[*]		
pΔ222, A197 cut & ligated w/ oligodeoxy- nucleotide pools:	<u>GGT ACG</u> TCA ATG GCA TCT CCG CAC GTT <u>GCA</u> GGA GCG-3'		
	CCA TGG AGT TAC CGT AGA GGC GTG CAA CGT CCT CGC-5'		
	<i>KpnI</i>	<i>PstI</i> destroyed	

Oligodeoxynucleotide pools synthesized with 2% contaminating nucleotides in each cycle to give
 -15% of pool with 0 mutations, -28% of pool with single mutations, and
 -57% of pool with 2 or more mutations, according to the general formula $f = \frac{\mu^n}{n!} e^{-\mu}$.

FIG.—35

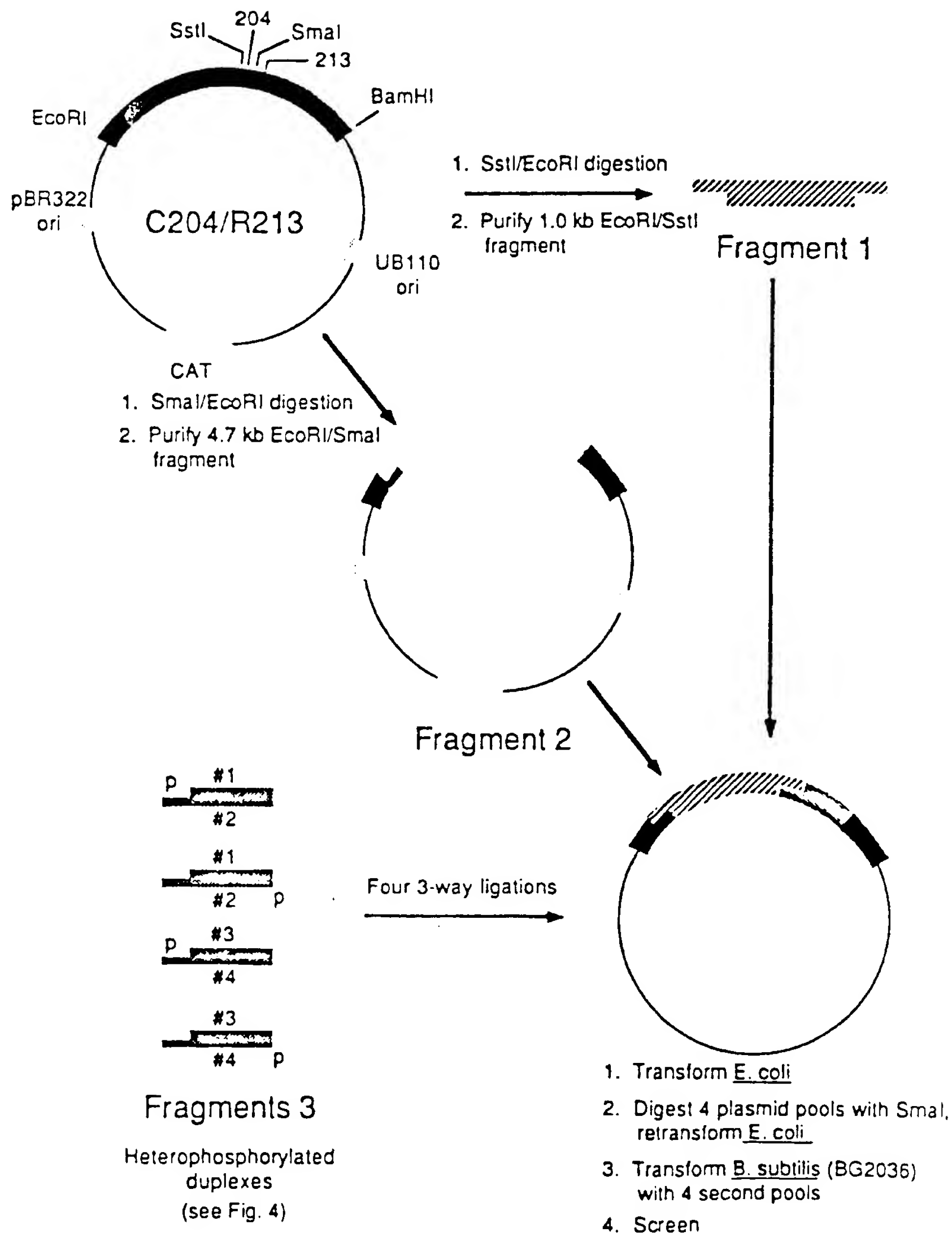


FIG.—36

	195	200	204	210	213																															
Wild type A.A.:	Glu	Leu	Asp	Val	Met	Ala	Pro	Gly	Val	Ser	Ile	Glu	Ser	Thr	Leu	Pro	Gly	Asn	Lys																	
Wild type DNA:	5'-GAG CTT GAT GTC ATG GCA CCT GGC GTA TCT ATC CAA AGC ACG CTT CCT GGA AAC AAA-3'																																			
	3'-CTC GAA CTA CAG TAC CGT GGA CCG CAT AGA TAG GTT TCG TGC GAA GGA CCT TTG TTT-5'																																			
C204/R213 DNA:	5'-GAG CTC GAT GTC ATG GCA CCT GGC GTA TGT ATC CAA AGC ACG CTT CCC GGG AAC AGA-3'																																			
	3'-CTC GAG CTA CAG TAC CGT GGA CCG CAT ACA TAG GTT TCG TGC GAA GGA CCC TTG TCT-5'																																			
	SsuI																SmaI																			
C204/R213 cut with SsuI and SmaI:	5'-GAG CT																GGG AAC AGA-3'																			
	3'-C																CCC TTG TCT-5'																			
C204/R213 cut and ligated with oligo-deoxynucleotide pools:	5'-GAG CTC GAT CTC ATG GCA CCT GGC GTA																ATC CAG TCG ACG CTT CCT GGG AAC AGA-3'																			
	3'-CTC GAG CTA CAG TAC CGT GGA CCG CAT																TAG GTC AGC TGC GAA GGA CCC TTG TCT-5'																			
	SsuI																SmaI																			
	<div style="display: flex; justify-content: space-around; align-items: center;"><div style="text-align: center;">$\begin{matrix} \text{W, R, R, or G} \leftarrow \text{NGG or} \\ \text{Stop, Y, H, Q, N, K, D or E} \leftarrow \end{matrix}$</div><div style="text-align: center;">$\begin{matrix} \text{NCC} \rightarrow \text{S, P, T or A} \\ \text{[G]AN} \rightarrow \text{L, F, I, V or M} \end{matrix}$</div></div>																																			

FIG.—37